DOCUMENT RESUME

ED 321 515 FL 017 668

AUTHOR Derbyshire, Desmond C., Ed.
TITLE 1988 Work Papers of the Summer Institute of Linguistics, University of North Dakota Session. Volume 32.
INSTITUTION Summer Inst. of Linguistics, Grand Forks, N. Dak.
PUB DATE 88
NOTE 220p.
AVAILABLE FROM SIL/UND Work Papers, c/o International Linguistic Center-Bookroom, 7500 West Camp Wisdom Road, Dallas, TX 75236.
PUB TYPE Collected Works - General (020) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC09 Plus Postage.
DESCRIPTORS Case (Grammar); Foreign Countries; *Language Patterns; Language Research; *Linguistic Theory; *Morphology (Languages); Phonology; *Portuguese; *Pragmatics; Semantics; Sentence Structure; Uncommonly Taught Languages
IDENTIFIERS *Anaphora; Brazil; Clitics; Mixtec; Tabaru; Word Order; Yagua

ABSTRACT

Six working papers in linguistics are presented. "The Semantics and Morphology of Mixtec Mood and Aspect" (J. Albert Bickford and Stephen A. Marlett) presents a relatively detailed description of the following grammatical categories in three Mixtec languages: Santiago Nuyoo, Santo Domingo Nuxaa, and Municipio of Tezoatlan. "Nonsyllabic Vocoids" (John M. Clifton) examples a number of New Guinean languages in an attempt to show that previous analysis of the relationship between syllabic i u and nonsyllabic y w is too simplistic and that when determining the markedness of a given segment, it is necessary to examine both paradigmatic and syntagmatic relationships. "Pragmatics and Grammar: Motivation and Control" (Robert A. Dooley) discusses pragmatic phenomena encoded in the syntax. "Anaphoric Indexes and Inalienable Possession in Brazilian Portuguese" (Daniel L. Everett) draws attention to some previously unnoticed facts about inalienably possessed noun phrases, which, it is argued, differ from both pronouns and anaphors. "Clitics, Case, and Word Order in Yagua" (Daniel L. Everett) argues that the relationships between two clitic types and their effect on word order present some problems for theories of syntax. "Tabaru Phonology and Morphology" (Edward A. Kotynski) examines recent phonological theory. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED321515

WORK PAPERS

VOLUME XXXII

1988

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

R. Dooley

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

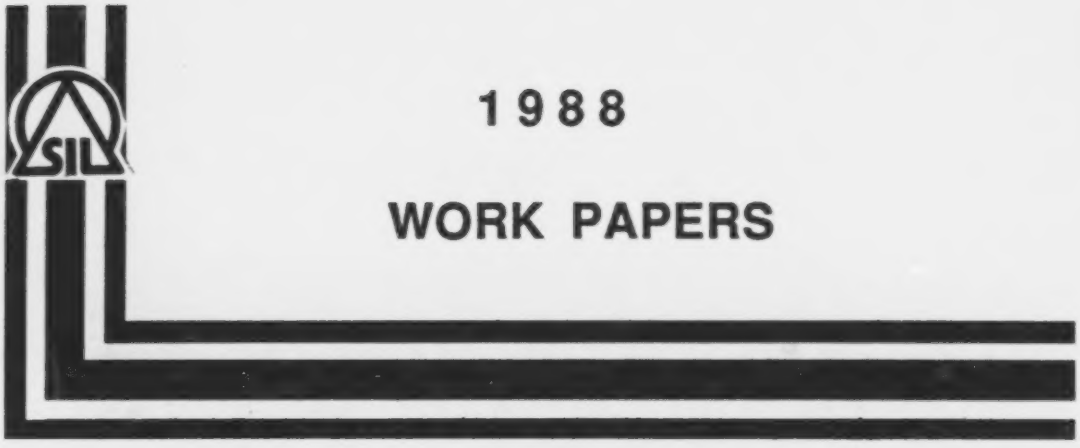
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

UND

SUMMER INSTITUTE OF LINGUISTICS
UNIVERSITY OF NORTH DAKOTA SESSION

FL017668

BEST COPY AVAILABLE



1988

WORK PAPERS

OF THE

SUMMER INSTITUTE OF LINGUISTICS

University of North Dakota

Session

Volume 32

Editor: Desmond C. Derbyshire

These are working papers and should not be cited without reference to their preliminary nature.

Orders and correspondence should be sent to:

SIL/UND Work Papers
c/o International Linguistic Center-Bookroom
7500 West Camp Wisdom Road
Dallas, TX 75236

PREFACE

Most of the papers in this volume present provisional results of ongoing research by five members of the SIL-UND faculty: Albert Bickford, John Clifton, Bob Dooley, Dan Everett and Steve Marlett. The sixth paper by Ed Kotynski is a slightly revised version of the thesis he presented to the UND Graduate School as part of the requirements for the M.A. (Linguistics) degree which he received in December 1987.

I would like to thank Betty Brown for the substantial part she played in the production of the volume, both in editing and in the composition.

D.C.D.

TABLE OF CONTENTS

Preface iii

The Semantics and Morphology of Mixtec Mood and Aspect 1
J. Albert Bickford and Stephen A. Marlett

Nonsyllabic Vocoids 41
John M. Clifton

Pragmatics and Grammar: Motivation and Control 59
Robert A. Dooley

Anaphoric Indices and Inalienable Possession in Brazilian Portuguese 87
Daniel L. Everett

Clitics, Case, and Word Order in Yagua 93
Daniel L. Everett

Tabaru Phonology and Morphology 143
Edward A. Kotynski

THE SEMANTICS AND MORPHOLOGY OF MIXTEC MOOD AND ASPECT

J. Albert Bickford
Stephen A. Marlett

---

1 Introduction
2 Grammatical categories of Mood and Aspect
  2.1 Aspectual distinctions within Realis Mood
  2.2 Realis vs. Irrealis
  2.3 On the nature of Mixtec 'Aspects'
3 The formal expression of Mood and Aspect
  3.1 Irregular consonant-initial roots
  3.2 Vowel-initial roots
  3.3 Regular consonant-initial roots
  3.4 Perfective and Imperfective aspect

---

1 Introduction

This study presents a relatively detailed description of the grammatical categories of mood and aspect in three Mixtec languages.[1] Our goals include morphology in a broad sense, covering both meaning and form. Our focus is a grammatical category known in Mixtec studies as 'aspect', which distinguishes three forms of every verb, generally called completive, continuative, and potential. We present a view of the system that differs somewhat from previous work, suggesting that the system is primarily modal, not aspectual in character. Partly, this is just a matter of how terms like 'mood' and 'aspect' are defined, but there is a more fundamental difference. We would like to suggest that the system is best analyzed in terms of two binary divisions, one between realis and irrealis mood, and a second within realis mood between perfective and imperfective aspect. We believe this leads to an insightful understanding of the meaning and usage of these categories, and also of their formal expression, inasmuch as it mirrors their semantic structure.

In the process, we are also able to give a first description of this system in three previously unstudied Mixtec languages. Our perspective is a comparative one, concentrating on those characteristics that are common to all three languages, and especially those which are found throughout Mixtec. Nevertheless, we have not attempted to survey the entire family; readers interested in a broader perspective should consult Bradley and Hollenbach (to appear). Also, we have not systematically explored uses of the three forms in complement clauses with different matrix verbs, nor have we been able to work out all the

details of tone and tone sandhi, since the tonal systems are as yet unanalyzed. In the future, we hope to address these deficiencies.

The three languages that we discuss are listed in (1), together with the names of our colleagues who specialize in each language, and who have collaborated closely with us.[2]

(1) Santiago Nuyóo (Nuy) - Larry Harris, Mary Harris
Santo Domingo Nuxáa (Nux) - Susan Hugghins, Inga McKendry
Municipio of Tezoatlán (Tez) - John L. Williams, Judith Williams

Indeed, the whole project has been a group effort, and could not even have been begun without the contributions of everyone. Our responsibility as authors has consisted in directing and coordinating the work of the others, and in producing the present analysis and discussion. Bickford is primarily responsible for the material in section 2 on the meanings and uses of the elements in the system, while Marlett is primarily responsible for the material in section 3 on their formal expression.

2 Grammatical categories of Mood and Aspect

Within the verbal paradigm, three basic forms are generally recognized under the labels Completive, Continuative, and Potential.[3] As we shall see below, it is initially tempting to analyze these as past, present, and future tense, respectively, since these represent their most common uses. However, this characterization is too narrow, as is well recognized by the conventional description of these three forms as Aspects, not Tenses.

However, the 'Aspects' of Mixtec differ in important ways from aspects of other languages. The traditional terminology adopts 'Aspect' as a general cover term for a system of verbal inflection which includes kind of action and possibly other factors in addition to time of action.

In what follows, we follow Comrie 1976 and Chung and Timberlake 1985 in using the term 'aspect' in a more restricted and precise sense which is more useful for cross-linguistic comparison; aspect in this sense is a grammatical category which refers only to the relationship of a situation[4] to some time interval selected by the speaker (Chung and Timberlake 1985:213-4). To put it another way, aspect refers to the internal temporal structure of a situation, rather than its location in time. This narrower sense includes concepts such as perfective, imperfective, habitual, and progressive. These terms are defined by Comrie (1976) in largely language-independent terms, yet they also work very well for characterizing the Mixtec system. We have therefore used them to facilitate comparison of Mixtec with other languages; we will give precise characterizations of them later.

Our primary difference with the traditional understanding of the Mixtec 'Aspects' is not terminological, however. We claim that the system does not involve a simple three-way distinction. Rather, there is a primary division of mood, between realis (actual) and irrealis (non-actual). The single Irrealis (= Potential) form stands in contrast to two realis forms, which are distinguished from each other as Perfective (= Completive) and Imperfective (= Continuative) aspect. A small class of motion verbs makes a further aspectual distinction between two types of imperfectivity, Habitual and Progressive. The structure of the whole system is represented in (2).

(2) Mood: Realis vs. Irrealis

Aspect: Perfective vs. Imperfective

Habitual vs. Progressive

In this view, then, aspectual distinctions are only relevant for the realis forms, while all forms include modal elements in their meanings. In particular, any precise characterization of the semantic structure of the Perfective, Imperfective, Habitual, and Progressive must recognize that realis mood is an important component of their meaning. This point will be demonstrated in detail below.

It will be necessary first to present the aspectual distinctions (section 2.1), since a familiarity with them is needed before we can proceed to the more fundamental distinction between them and the Irrealis (section 2.2). Section 2.3 summarizes the semantic discussion and pulls together in one place the reasons for considering the system as being primarily modal, with aspect secondary, as opposed to other reasonable alternatives.

2.1 Aspectual distinctions within Realis mood

As indicated above, a naive analysis of the two principal realis forms would claim that they indicated past and present tense. Indeed, this hypothesis would cover most instances of their use, especially in ordinary conversation, and is reflected in the labels given to them in many practical grammars of Mixtec languages written for nonlinguists, such as Alexander 1980:22ff, Dyk and Stoudt 1965:125ff, North and Shields 1976, Pensinger 1974:142ff, and Stark Campbell, Johnson Peterson, and Lorenzo Cruz 1986:161ff (although several of these sources also point out that temporal labels do not adequately characterize the meanings of these forms).

The Perfective is used for isolated past events, or for a series of past events in narrative discourse. It is normally indicated by a clitic **n(i^{a})** preceding what we call the *Realis* form of the verb, together with as-yet unanalyzed tonal changes. (Part of our analysis

involves positing a Realis prefix within this form which distinguishes it from the Irrealis form. Throughout section 2, we do not isolate this prefix, leaving such matters to section 3.)

(3) iku ǹ ku?wi =nì?. (Nuy)
yesterday Perf sick 1s
I was sick yesterday.

(4) ǹ sà?ánoo =ránì? nùu?ⁿ yuuⁿ
Perf clear 1pe face ground
We (excl.) cleared the field

xee ǹ tee =ránì? nùnì?. (Nuy)
and Perf plant 1pe corn
and we planted the corn.

(5) iku niⁿ seeⁿ =ða waaⁿ kʷi?i. (Nux)
yesterday Perf buy 1r much fruit
Yesterday I bought a lot of fruit.

(6) niⁿ teku waaⁿ =niⁿ te niⁿ kee =yaⁿ ke?eⁿ. (Nux)
Perf listen mother 1 and Perf go.out 3r outside
My mother heard, and she came outside.

(7) kúú nìⁿ saà =ta,
and Perf arrive 3s,
And he arrived,

kúú nìⁿ nàní?ìⁿ tuku =ta ta θito ilo. (Tez)
and Perf find again 3s 3s uncle rabbit
and he again found uncle rabbit.

The Imperfective is used to refer to present situations, specifically events in progress at the moment of speech.

(8) té?néni =nì? čìxi =nì?. (Nuy)
Impf/hurt 1s stomach 1s
My stomach hurts.

(9) tee =ða waaⁿ tu?uⁿ. (Nux)
Impf/write 1r much words
I am writing many words.

(10) sásá?aⁿ =naⁿ. (Tez)
Impf/eat 3s
He/she is eating.

However, other data shows that the distinction is aspectual, not temporal. When Perfective and Imperfective forms occur together, the Perfective refers to a span of time which is completely included in the

span referred to by the Imperfective. In this case, the Imperfective, as well as the Perfective, is used for past time.

(11) iku ká?wi wí?í ini =yaⁿ
yesterday Impf/worry much inside 3f
Yesterday, she was very worried

sá ǹ nenta =rè?. (Nuy)
while Perf arrive 3m
when he arrived.

(12) weeⁿ=yi seki nuči
FOC=3m Impf/plant beans
He was planting beans

te niⁿ see yaniⁿ =yi. (Nux)
and Perf arrive brother 3s
when his brother arrived.

(13) síonooⁿ ta θito ilo sásá?aⁿ =ta yúkù;
Impf/walk 3s uncle rabbit Impf/eat 3s mountain
Uncle rabbit was walking along eating in the country;

iiⁿ kuù nàⁿkitá?aⁿ =ta sí?íⁿ ta θito nikʷè?í. (Tez)
one day meet 3s with 3s uncle coyote
one day he met up with uncle coyote.

As can be seen from these examples, Perfective forms are frequently used for the basic plot line of a narrative, while Imperfective forms are used for background material. This is a consequence of the fact that most background situations occur over a longer span of time than the events that make up the plot line, and include them.

What is common to all uses of the Perfective is that it presents a situation as a complete whole, whose actual internal complexity (if any) is not relevant to the discourse. Thus, it is used for most past situations, which are viewed 'from the outside' (at the moment of speech); the sequence of subevents that make them up is in such cases not of interest, only the event taken as a whole.

In contrast, the Imperfective presents a situation whose internal complexity (e.g., the series of states that make up a process) is relevant to the discourse, and views situations 'from the inside out'.[5] Note that this is not a matter of the situation itself, but rather of how the speaker chooses to represent it and its relationship to other situations. There are two possible ways of understanding how a speaker's conceptualization of a situation as Imperfective involves an 'inside out' perspective.

From one point of view, the internal structure of a situation is relevant because some portion, but only a portion, of its time span coincides with some other significant location in time, or reference point. In its use for present situations, the reference point is the moment of speech. In its use in contrast to the Perfective, the reference point is the time of the Perfective situation.

From the other point of view, the internal structure is relevant because only the first portion of the situation is realis with respect to the reference point (either the present moment or a Perfective situation). As we discuss in the next section, one of the senses of the realis forms is that of relative nonfuture. When the entire situation is realis with respect to the reference point, a Perfective form is used; when only a portion is realis, Imperfective is used instead.

In addition to these uses in past and present, both Perfective and Imperfective forms can be used for future situations, although such uses involve complexities that are best postponed until the discussion of the Irrealis. Thus, the association of Perfective with past and Imperfective with present is not due to anything inherent in their meaning, but is simply a matter of customary usage. Both forms can be used, with the same essential aspectual meanings, for situations without regard to their location in time.[6]

In much work on Mixtec, such as Daly 1973:17ff, and the papers in Bradley and Hollenbach (to appear), the Perfective is called Completive, and the Imperfective is called Continuative. The term 'Continuative' should not be confused with the term 'continuous', which is generally used to represent that portion of imperfectivity which is not habitual (Comrie 1976:26, Crystal 1985:247).[7] However, the Mixtec Imperfective (= Continuative) is used for habituality as well as continuousness.

(14) sá?á ránì? wiko ta *Semana Santa*. (Nuy)
Impf/do 1pe fiesta every week holy
We have a fiesta every (year during) Holy Week.

(15) setu =yi niaani^{n} nuu. (Nux)
Impf/work 3m all day
He works every day.

(16) niθaa kuù kúčóo^{n} =na^{n}. (Tez)
every day Impf/work 3r
He/she works every day.

These examples are consistent with the characterization of the Imperfective given earlier, provided that we understand habitual situations as consisting of a macrosituation formed of several component situations. When this macrosituation includes the present moment, it is viewed from the inside out, just as with simple situations in progress in the present.

However, that portion of the macrosituation highlighted by the Imperfective need not be the present moment; it can also be the time referred to by an adverbial phrase, such as 'every May' or 'every year' in (17).

(17) ta yòo? Mayú?/kʷlya?
every month May /year
Every month of May / every year

sá?ánoo =ránl? nùu?ⁿ yu?uⁿ
Impf/clear 1pe face ground
we (excl.) clear the fields

xee táxi =rán? nùni?. (Nuy)
and Impf/plant 1pe corn
and plant the corn.

To show the unity of meaning of the Imperfective, the same verb form can be used with a nonhabitual reading, which can be forced with the proper adverbial.

(18) wita táxi =ránl? nùnl?. (Nuy)
today Impf/plant 1pe corn
We (excl.) are planting the corn today.

Similarly, the Imperfective can be used for habituality in the past, at least in Nuyóo. The following examples are from a discussion of former marriage customs:

(19) xá?a =ráa ntewísí?... (Nuy)
Impf/give 3p soft.drink
They (the bride's family) used to give soft drinks (to the suitor's family)...

(20) ...kʷáno?o =ráa náke?eⁿ =ráa ya?aⁿ. (Nuy)
Impf/go 3p Impf/get 3p woman
...they used to go get the bride (after the wedding feast was over).

As a further clarification of the meaning of the Imperfective, we must add that mere habituality of a situation is not sufficient to place a verb in the Imperfective.

(21) sá ǹ nee =nl? yuuⁿ nuyóó
while Perf live 1s town Nuyóo
While I lived in Nuyóo,

xee ǹ yaxi =nl? we?i alwétó?. (Nuy)
and Perf eat 1s house Alberto
I ate (used to eat) at Alberto's house.

Since the time span of the macrosituation of habitually eating coincides exactly with the time span of living in Nuyóo, the internal structure of both situations is irrelevant to the discourse. Both can be treated as complete wholes, in the Perfective. In other words, what makes most habitual situations Imperfective is not habituality *per se*, but the coincidence of some other time span (such as the moment of speech) with a portion of the habitual macrosituation, imposing an 'inside out' perspective on it.

These considerations result in interesting pairs such as (11), repeated here as (22), and (23).

(22) iku ká?wi wí?í ini =yaⁿ
yesterday Impf/worry much inside 3f
Yesterday, she was very worried

sá ǹ nenta =rè?. (Nuy)
while Perf arrive 3m
when he arrived.

(23) ǹ kà?wi wí?í ini =yaⁿ
Perf worry much inside 3f
She was very worried

yà ntu ǹ kenta =rè? iku. (Nuy)
that Neg Perf arrive 3m yesterday
that he didn't arrive yesterday.

In (22), the worrying forms a backdrop to a single action of arriving, so 'worry' is in the Imperfective. However, in the very similar (23), the time span of worrying is identical to that of the non-arrival, so Perfective forms are used in both clauses.

A handful of motion verbs meaning 'go' and 'come' in each language make a distinction between two types of imperfectivity, Habitual (sometimes called 'Iterative') and truly continuous, or Progressive.[8] This extra aspectual distinction was first described in Texmelucan Zapotec, another Otomanguean language, by Speck and Pickett (1976).[9] Subsequently, Macaulay (1985) has argued that it is also present in two Mixtec languages, Chalcatongo and Diuxi, and that positing three aspects for verbs of motion solves several problems in an earlier analysis of Diuxi (Kuiper and Merrifield 1975), which only recognized two. In

particular, it allows a straightforward account of the fact that Perfective forms generally represent a round trip, whereas Progressive forms are used for motion in one direction only.

We illustrate with examples from Nuxáa and Tezoatlán, for which this analysis also seems to be correct. (The analysis of motion verbs in Nuyóo has not been completed.)

(24) kʷa?aⁿ =yi nuuⁿ nuwa. (Nux)
Prog/go 3m Oaxaca.City
He went to Oaxaca (and hasn't returned).

(25) nìⁿ sà?àⁿ nà?í wíiⁿ =i ninòⁿ káa,
Perf go poor self 1s below there
I went down there, poor me,

níⁿ sà nè?é =í noòⁿ ití =ì. (Tez)
Perf Dur see 1s face cornfield 1s
and looked at my corn field.
(spoken after returning home)

Under an analysis that posits the extra Habitual/Progressive distinction, all such verbs refer to motion away from a point and subsequent return to the same point. (The differences between the different 'go' and 'come' verbs depend on the location of this point with respect to both the place of speech ('here') and the subject's 'home base', or customary location.) Perfective forms, representing the action as a complete whole, naturally refer to round trips.[10] Progressive forms, being imperfective, refer to a trip in progress at the moment of speech, that is, one that has begun but has not been completed by return to the point of origin. They can also refer to a trip that is begun, but never completed, or whose completion is irrelevant to the discourse. That is, one-way trips are generally expressed in the Progressive, even trips mentioned as one of a series of events in the past, a context that would require the Perfective for other verbs.

(26) niⁿ nakʷi?nuⁿ =niⁿ lona paa =niⁿ. niⁿ sika =niⁿ.
Perf put.on 1 cloak father 1 Perf walk 1
I put on my father's cloak. I walked.

kʷa?aⁿ =niⁿ kʷeeniⁿ nuⁿ nee kolo. (Nux)
Prog/go 1 slowly place be turkey
I went slowly to the place where the turkey was.

(27) θa nìⁿ kàki =tí kʷa?àⁿ-wa =tí. (Tez)
then Perf escape 3a Prog/go 3a
Then he escaped, and went away.

(28) θá nlⁿ kee =ta kʷa?àⁿ =ta
then Perf leave 3m Prog/go 3m
Then they left, they went,

nlⁿ saà =ta iniⁿ wìniⁿ yoóⁿ. (Tez)
Perf arrive 3m in pond there
and they arrived at a pond.

This results in an apparent anomaly in translation between Mixtec and a language (such as English) in which motion verbs are unidirectional. The Progressive form in these examples, even though it is imperfective, is most appropriately translated with a perfective form. Yet, the anomaly disappears when one takes into account the differences in the meanings of the verbs in each language.

Habitual forms of these verbs hold no surprises under this analysis; they simply indicate that a round trip was repeated habitually over a period of time. Note the use of the Habitual alongside an ordinary verb in the Imperfective.

(29) kúú niθaá kuù wa sá?àⁿ =ti seí =ti =áⁿ. (Tez)
and every day Intens Hab/go 3a Impf/eat 3a 3n
And every day he used to go and eat them.

This sentence occurs in a story immediately preceding a sequence of verbs in the Perfective, and forms the background against which the sequence is presented. This shows that the Habitual of motion verbs is imperfective in the same way as Imperfective forms of other verbs.

In general, the Habitual form is morphologically identical to what one would expect as the Imperfective form of these verbs. To the extent that inflection is regular, Imperfective forms consist of the bare Realis form, analyzed in more detail in section 3. That is, the segmental material of these forms is identical to what is found in the Perfective, minus the proclitic **n(iⁿ)**. The same is true of Habitual forms for motion verbs; for example, consider the motion verbs of Tezoatlán in (30).

(30)

	Perf	Hab	Prog	Irrealis
Go from base	nl sà?àⁿ	sá?àⁿ	kʷà?àⁿ	kù?ùⁿ
Go to base	nl nò?òⁿ	nó?òⁿ	kʷànó?òⁿ	no?òⁿ
Come from base	nl kisi	kísi	wàsi	kisi
Come to base	nl nisi	nisi	kònii/ kʷèⁿnisi	nisi

As can also be seen in (30), the Progressive is always morphologically distinct from the Imperfective form of other verbs.

It would appear, then, that the Habitual *is* the Imperfective form of motion verbs. That is, the bare Realis form is generally used with a broad imperfective meaning, but for those verbs which have a special Progressive form, the meaning of this form is narrowed to cover just habituality.[11] Or, to put it another way, the bare Realis form is the general purpose means of indicating imperfectivity, although other, more specific forms (such as the Progressive) may encroach on part of its semantic territory at times, leaving it with a narrower meaning. The meaning of the bare Realis form, which we have glossed either Imperfective or Habitual, is thus 'that portion of imperfectivity which is not covered by a competing form (if any)'.

To sum up, then, the Perfective ignores the internal structure of a situation, and presents it as an unanalyzed whole. The Imperfective, on the other hand, accords prominence to the internal structure of a situation, and in particular, to some subportion of its time span. A few motion verbs make finer distinctions within imperfective aspect, and have special Progressive forms; the Imperfective form of these verbs consequently refers only to habituality.

2.2 Realis vs. Irrealis

We turn now to the Irrealis form, often called 'Potential' in other Mixtec studies. For many verbs, the Irrealis is identical (except for tone) to the Realis form, but for others, the two forms are segmentally as well as tonally distinct. Section 3 discusses the differences in more detail.

The Irrealis designates situations that are nonactual, as opposed to the Perfective and Imperfective, which are used only for actual, or *realis* events. In Mixtec languages, there are three basic ways that situations can be non-actual. They can represent an imposition of the speaker's desires on the world around him, as in a command or wish; they can be statements about situations which are not yet actual (because they are future);[12] or they can be statements about hypothetical (or even contrafactual) situations, as in conditional clauses.

The simplest type of command uses the Irrealis without any additional imperative morphology, and generally lacks an overt subject noun phrase.[13]

(31) ko?o waaⁿ. (Nuy)
Irr/drink it
Drink it!

(32) ko?o rán. (Nuy)
Irr/drink you(pl.)
(Y'all) drink!

(33) kasi δita. (Nux)
Irr/eat tortilla
Have a tortilla!

(34) koo. (Tez)
Irr/sit
Sit down!

(35) kata =ní^{n}. (Tez)
Irr/dig 2sr
Dig!

All of the above examples use the bare Irrealis form, without additional modal morphemes, and express a strong, direct, positive command. Positive commands that are more polite, and all negative commands, require special particles or special constructions in addition to the irrealis.

The clearest cases of this are in Nuyóo, where polite commands are not expressed syntactically as commands at all. For example, the most polite way to express a request is to phrase it as a yes-no question in the Irrealis, and to include the morpheme šá, which elsewhere seems to express an extreme degree of an abstract quality.

(36) a kuwí kuwí^{n} =nú **šá** sè?yá =nì?. (Nuy)
Q Irr/able Irr/care.for 2 **extreme** child 1
Would you be able to care for my child, please?

However, the most common way of forming a polite command in Mixtec languages is apparently to use the Irrealis preceded by the particle **na(^{n})**, which is variously described as 'hortatory' (Bradley 1970:41, Bradley and Hollenbach, to appear), 'subjunctive' (Alexander 1980:35, Macaulay 1987:116-21), 'permissive imperative' (Stark Campbell, Johnson Peterson, and Lorenzo Cruz 1986:164), the 'prefix of permission' (Pensinger (1974:141), or a 'particle of constraint' (Daly 1973:17). Macaulay (1987) suggests the gloss 'Optative' (following Bybee (1985:166)), at least for most dialects. This seems to capture the core of its meaning in Tezoatlán, in that it expresses a wish or desire of the speaker.

(37) ná^{n} kuθì wà?a =ní^{n}. (Tez)
Opt Irr/sleep good 2r
May you sleep well.

(38) θá ná^{n} saà ilo,
then Opt Irr/arrive rabbit
Then the rabbit will hopefully arrive,

θá ná^{n} tìi^{n} =tí kana?à =tì šaà^{n}. (Tez)
then Opt Irr/grab 3a Irr/stick 3a there
and will (hopefully) grab and be stuck there.

In Nuxáa, **na^{n}** almost always occurs when the subject is first person. When the subject is first person singular, it expresses what the speaker intends to do.

(39) na^{n} čik^{w}a?a =δa =si; tasi =ni nuu^{n} kee =si. (Nux)
Opt Irr/measure 1r 3 Irr/give 1 Loc place 3
I intend to measure it; give me something to put it in.

When the subject is first person plural, the expressed desire can be understood, by implicature, as a polite command.

(40) na^{n} ki?i^{n} =no. (Nux)
Opt Irr/go 1
Let's go! (= May we go!)

In Tezoatlán, on the other hand, **na^{n}** is always optative, without ever having the force of a command.

Negative commands use various negative particles in place of **na^{n}**. In Nuyóo, a negative statement in the Irrealis can be used with the force of a command.

(41) n- kó?o =ń waa^{n}. (Nuy)
Neg Irr/drink 2 3n
Don't drink it! / You are not going to drink it.

More commonly, however, a negative command consists of **(ko)to** 'negative imperative' followed by, oddly enough, the Imperfective form of the verb.

(42) koto táhi =ń sè?ya =nì?. (Nuy)
Neg/Imp Impf/give.me 2 child 1
Don't give me my child!

This use of the Imperfective in what appears to be an irrealis context remains unexplained. Inasmuch as language-particular categories rarely coincide completely with universal ones, perhaps no explanation is necessary; this is apparently just a quirk in the system.

Nuxáa uses the Irrealis with the ordinary negative particle **ya^{n}** plus a special negative imperative particle **si**.

(43) yaⁿ si kasi =ni =si. (Nux)
Neg Neg/Imp Irr/eat 2r 3
Don't eat it!

Tezoatlán uses the Irrealis with **náθá?ani** 'negative imperative' for direct commands, and the milder (k)os(a) for giving advice or suggestions.

(44) náθá?ani ka?ⁿ =òⁿ. (Tez)
NegImp Irr/speak 2
Don't speak!

(45) kòs(à) ká?àⁿ =níⁿ. (Tez)
NegAdv Irr/speak 2r
You shouldn't speak.

In summary, then, with the exception of one construction in Nuyóo, Irrealis forms are used for expressing wishes and commands.

The second major use of the Irrealis is for future situations, those that are not yet actual.

(46) nénù? ke?eⁿ ránl? aros kaxi? ránl?. (Nuy)
where Irr/buy 1pe rice Irr/eat 1pe
Where will we (excl.) buy rice to eat?

(47) ki?iⁿ =δa kaδi =δa ya?a. (Nux)
Irr/go 1r Irr/grind 1r chili
I am going to go and grind chili.

(48) ko?òⁿ =ì kʷiiⁿ =i yà?aⁿ kasá?aⁿ =i. (Tez)
Irr/go 1s Irr/buy 1s thing Irr/eat 1s
I am going to buy things to eat.

A sentence with the appropriate subject can thus function equally as a statement about the future or as a command. The pragmatic context of such an utterance must be considered to determine if it is a command or statement; the semantic structure of Mixtec does not distinguish the two. Looking at translations into languages like English, which formally distinguish commands and future statements, one might be tempted to say that the Irrealis is ambiguous between the two readings. However, this is not true ambiguity, but simply vagueness. Both commands and future time involve non-actuality; this is all that the Irrealis expresses, without making any finer distinctions.

In subordinate clauses, the Irrealis expresses situations subsequent to that expressed by the matrix verb, i.e. situations that are relative future with respect to the main verb.

(49) íyó kɨ̀?ɨ́? =nì? *tienda*
Impf/exist Irr/go 1 store
I was going to go to the store
(lit., It was existing that I would go to the store),

soo ǹ kinòo =nì? we?i. (Nuy)
but Perf stay 1 house
but I stayed home.

(50) satíyúⁿ =ó nèè? kʸaa =rè?. (Nuy)
Irr/work 1pi until Irr/arrive 3m
We will work until he arrives.

(51) a. θá nìⁿ ka?aⁿ =tà θito nikʷè?í keí-yáⁿ =tà,
then Perf think 3m uncle coyote Irr/eat-DO 3m
Uncle Coyote thought he would eat him (rabbit),

tíθo nìⁿ θanà?í-yáⁿ =tà θito ilo. (Tez)
but Perf trick-DO 3m uncle rabbit
but Uncle Rabbit tricked him.

b. niⁿ kaa =tí keí-yáⁿ =tí
Perf say 3a Irr/eat-DO 3a
He said that he would eat him,

kúú witiwiti-o nìⁿ seí-yáⁿ =tí. (Tez)
and right.away Perf eat-DO 3a
and right away he did eat him.

As the examples in (51) show, it is not the actuality of the event from the speaker's point of view at the moment of speech that is relevant in subordinate clauses. The actuality of these situations is judged from the time of the matrix, so that situations subsequent to that represented by the matrix verb are encoded as Irrealis.[14]

Similarly, embedded commands are Irrealis, in that the imperative modality is judged with respect to the matrix subject, not the speaker. In this example, the speaker did in fact go to tend the oxen; the proposition expressed by the embedded command was thus realis with respect to the speaker at the moment of speech. Yet, the Irrealis is used, because the embedded proposition is irrealis with respect to the matrix clause.

(52) niⁿ ka?aⁿ δito =niⁿ
Perf tell uncle 1
My uncle told me

ha ki?iⁿ =niⁿ kunʸa?a =niⁿ ɲjikutu =yaⁿ. (Nux)
Comp Irr/go 1 Irr/tend 1 oxen 3r
to go look after his oxen.

A special case of this use for relative future is that all clauses expressing purpose are Irrealis. This can be seen in several examples above, especially when the main verb is a verb of motion; such verbs frequently appear with purpose clauses. However, all of the previous examples involve purpose relative to future situations, so that the purpose clause might also be claimed to use Irrealis because it itself is future with respect to the present moment. The following examples show that Irrealis is also used to express purpose relative to present and past situations. The purpose clauses are in brackets; note that both adverbial and relative clauses of purpose use Irrealis.

(53) tá kiwɨ́ ké?eⁿ =nì? yutu
every day Impf/buy 1 wood
Every day I buy wood

[(sáwa) sa?á we?i]. (Nuy)
(for.to) Irr/make house
to make a house.

(54) na órá? kéne =nú
what hour Impf/leave 2
What hour do you leave

[nasánoo =nú itu]? (Nuy)
Irr/clear 2 field
to clean the fields?

(55) kʷiniⁿ =yi ha δiko =yi nuku [či kʷeeⁿ
Impf/want 3m Comp Irr/sell 3m wood because Irr/buy
He wants to sell wood, so as to buy

yi yuku [ha ko?o aδi?i =yi]]. (Nux)
3m medicine Comp Irr/drink wife 3m
medicine for his wife to drink.

(56) káwànlóⁿ kasá?á yu?ù kenu =l yà?a [kasá?a =nu]
lack Irr/begin I Irr/prepare 1s thing Irr/eat 1pe
Just before I began to prepare things for us to eat,

kuu sà nlⁿ kàsáà wa =na. (Tez)
and Recent Perf arrive just 3
they arrived.

Interestingly, there is one type of future situation which is classified as realis. A subordinate clause will generally use realis forms if it is presupposed.[15] This is true whether the main clause is a command or statement.

(57) ni súkʷá?a =nl? sekuntáriá? inka kʷlya?
while Impf/study 1 secondary other year
While I study in the secondary school next year,

xee kɨ?ɨ? =nl? kaxi =nl? we?i násúčí =nl?. (Nuy)
and Irr/go 1 Irr/eat 1 house aunt 1
I will eat at my aunt's house.

(58) ki niⁿ nani?i =ða ha niⁿ sewa?a =ða kiði
when Perf finish 1 COMP Perf make 1 pot
When I finish (have finished) making the cooking pot,

ðate kawa?a =ða iiⁿ šoo. (Nux)[16]
then Irr/make 1r one griddle
I will make a griddle.

(59) tá kʷa?à =òⁿ
when Prog/go 2
When you go,

ta ki?i =oⁿ iiⁿ kiló níkà, kl?o ló?ò. (Tez)
then Irr/get 2 one kilo banana sister little
get a kilo of bananas, sister.

(60) tá nlⁿ ni?i nlⁿ siiⁿ =nu tatà,
when Perf finish Perf buy 1p medicine
When we finish (have finished) buying medicine,

θa ko?òⁿ =nu we?e ta Ninfa. (Tez)
then Irr/go 1p house Dim Ninfa
then we will go to Ninfa's house.

Apparently the presupposition that some future situation will occur is sufficient to classify it as actual, and thus a realis form is used instead of the Irrealis. It is thus necessary to amend the statement given earlier that Irrealis mood is used for future situations. To be

more precise, Irrealis includes only situations that are asserted in the future.

Note that the basic aspectual distinctions discussed in section 2.1 determine the choice of form here. Progressive with motion verbs is used for trips begun but not completed. Imperfective is used with other verbs for situations whose time span includes that of the main verb. Perfective is used for situations that can be treated as complete wholes, typically because they are terminated before the time of the main verb.[17]

The third major use of the Irrealis is for hypothetical situations, i.e., those that are discussed in general terms, with no claim or presupposition as to their actuality. Mostly, these occur in conditional clauses, in which the main clause serves as the consequent.

(61) ré ke?eⁿ =nì? tʸuúⁿ wita
if Irr/buy 1 chicken today
If I buy a chicken today,

xee kaxi =nì? waⁿ tee. (Nuy)
and Irr/eat 1 3Indef tomorrow
I will eat it tomorrow.

(62) te noⁿ naⁿ kasi =ða waaⁿ ha wiði
and if Opt Irr/eat 1r much Rel sweet
And if I eat a lot of sweets (as I would like to do),

te kunu?u =ða. (Nux)
and Irr/fat 1r
I will be fat.

(63) totá náⁿ kooⁿ θal,
if Opt Irr/fall rain
If it rains,

kúú nasi wa θá?ðⁿ káa. (Tez)
and Irr/get.wet Intens clothes those
the clothes over there will get wet.

Typically, the consequent is also hypothetical, and thus is also Irrealis. However, a hypothetical condition can also be paired with a realis consequent, as in a procedural discourse like a recipe.

(64) ré yà? kaxi? ntayú?úⁿ wí
if that Irr/eat now is
If it's to eat now,

xee wɪtá sá?á ráa waⁿ. (Nuy)
and soft Impf/make 3pIndef it
then they make it soft.

The Imperfective form in the consequent expresses what is habitually done under the circumstances expressed by the Irrealis form in the condition. Because of the habituality, the consequent can be construed as actual, and expressed with a realis form.

Closely related semantically to hypothetical situations are contrafactual ones, i.e., situations that are presupposed to be false. However, there are extra formal complications with contrafactual conditions in Mixtec. Surprisingly, many contrafactual conditions require realis forms, even though they refer to situations that are presupposed to be nonactual. The contrafactual consequents continue to be in the Irrealis.

(65) ríá? iyo šú?ù =nɪ?
if Impf/exist money 1
If I had the money,

soo kɨ?ɨ? =nɪ? siné? yakʷáá? wita. (Nuy)
but Irr/go 1 movie night today
I would go to the movie tonight.

(66) riá? ǹ kuneí Maximiliano ǹ kàntá?áⁿ
if Perf win Maximilian Perf fight
If Maximilian had won the war,

soo ka?niⁿ Juarez. (Nuy)
but Irr/kill Juarez.
he would have killed Juárez.

(67) totá níⁿ kóoⁿ θaì, kúú
if Neg/Perf fall rain and
If it had rained,

nasi wa θá?òⁿ yóóⁿ níⁿ kúú. (Tez)
Irr/get.wet Intens clothes those Neg/Perf happen
the clothes would have gotten wet, (but) it didn't happen.

In other contrafactual conditions, however, Irrealis forms are used. Such sentences are apparently identical to the hypothetical condition/consequent pairs discussed above.

(68) ré kuu sàwl yàne?eª wita
if Irr/fall rain morning today
If it had rained this morning,

saa naxi nùu?ª yu?uª. (Nuy)
while Irr/wet face ground
the ground would be wet (now).

The generalization, which holds in other examples that we have tested, seems to be that the form used in the condition indicates the time at which the contrafactuality of the consequent is evaluated. A Perfective form in the condition indicates that the consequent is contrafactual in the past, i.e., the situation would have happened in the past. An Imperfective form in the condition indicates that the consequent is contrafactual in the future. An Irrealis form in the condition indicates that the consequent is contrafactual in the present.

Although our understanding of these sentences is imperfect, it would seem that in this one corner of the grammar, the three basic mood/aspects are used in a way that is completely different from their ordinary uses. It is probably best simply to accept this as a quirk in the Mixtec modal system, without requiring further explanation for it.

Overall then, with a couple of idiosyncratic exceptions, the Irrealis is used for situations that are nonactual, such as wishes and commands, future situations, and hypothetical or contrafactual situations. Realis forms are used for actual situations. In main clauses, actuality is evaluated from the perspective of the speaker; in subordinate clauses, it is evaluated with respect to the subject and verb of the matrix clause.

2.3 On the nature of Mixtec 'Aspects'

The Mixtec system of 'Aspects', then, can be seen to be dominated primarily by the modal categories Realis and Irrealis. As we have seen, Perfective and Imperfective forms are used only for actual situations (with only minor, idiosyncratic exceptions); and Irrealis forms are used only for nonactual situations. A full characterization of the meaning of the first two forms must therefore include the modal specification that these two forms are realis, at least in their most general usage. On the other hand, the aspectual distinctions, such as perfectivity, are relevant only for the realis forms, not for the Irrealis. Mood is thus more important than aspect in this system, in that the modal specification of realis vs. irrealis must be included in the meaning of all three basic forms, whereas aspectual specifications need only be included for realis forms.

As we pointed out earlier, it is clear that the system is not one of absolute tense. Imperfective forms may be used for past as well as present situations, and both Perfective and Imperfective forms may be used for future situations, if they are presupposed.

It is not so obvious that the system could not be analyzed as one of relative tense, in which Perfective would be analyzed as Relative Past, Imperfective would be Relative Present, and Irrealis would be Relative Future. Many of the uses noted above could be accounted for in this way. For example, in cases where the Perfective is used with future reference in presupposed subordinate clauses, it is always relative past with respect to the situation in the main clause.

Such an analysis falls short in at least two ways. First, it doesn't account for the use of the Irrealis to indicate hypothetical situations, which are essentially timeless. Indeed, if this was a tense system, one might expect a (Relative) Present to be used for hypothetical situations, paralleling the facts in many true tense systems. The fact that one form is used both for future and hypothetical situations shows that it is broader in meaning than Relative Future; it is Irrealis. The relativity that is clearly present is a normal part of modal systems; that is, mood in subordinate clauses is normally evaluated with respect to the matrix subject and verb. (See Chung and Timberlake 1985:241ff.)

Second, an analysis of relative tense fares even more poorly when considering the uses of Perfective and Imperfective together in the same sentence. In all cases, it is true, the Imperfective indicates a situation that is present relative to the situation expressed with the Perfective. However, the Imperfective is not always in a subordinate clause, as would be expected in a system of relative tense. Worse, the Perfective form does not express a situation that is prior to the situation expressed in the Imperfective. Even though relativity enters into the modal distinction which is relevant for comparing the Perfective and Imperfective to the Irrealis, relativity is not important in the aspectual distinctions within the realis domain. Further, an analysis of relative tense would not provide any way of characterizing the difference between Progressive and Habitual forms of motion verbs; aspectual distinctions would be needed at least for these forms.

On the whole then, the system is best analyzed in terms of mood, with aspect secondary. It is not in any way a tense system, either absolute or relative.

3 The formal expression of mood and aspect

The analysis of the Mixtec verb system presented above finds some support in the segmental and suprasegmental facts. This support is not as strong as one might expect, and previous studies have not given much importance to the facts discussed below. Nevertheless, in this section we show the extent to which these facts can be taken in support of the analysis given above according to our present understanding of them.

At least some verbs in Mixtec languages with which we are familiar display the pattern shown in (69) with respect to the segmental alternations.

(69)	Imperfective	Perfective	Irrealis
	B+C	A B+C	D+C

These patterns can be seen in the forms of the verb 'walk' from Tezoatlán given in (70).

(70)	Imperfective	Perfective	Irrealis	
	sika	ni^{a} sika	kaka	'walk' (Tez)

The syllable ka occurs in all forms. However, the syllable si occurs in the Imperfective and Perfective forms and not in the Irrealis form.

We must be quick to point out that most verbs in Mixtec do not show such alternations. Nevertheless, we take these facts as supporting the view that the Imperfective and Perfective forms should be grouped together as realis forms. In fact, at least for a small class of verbs, there is an overt marker of realis mood and an overt marker of irrealis mood. We now turn to the analysis of mood morphology.

The basic division of Mixtec verbs, in our analysis, is along the following lines:

(71) a. Regular consonant-initial
b. Regular vowel-initial
c. Irregular consonant-initial

There are some irregular verbs which do not fit into these groups, which we mention below but of which we do not give an integrated account. As it turns out, the regular consonant-initial verbs are the most numerous across Mixtec, but they are also the verbs which typically reveal the least information about the segmental morphology of mood. The regular vowel-initial verbs preserve some of the segmental information, but it is actually the irregular consonant-initial verbs which give the most information.

The picture which emerges is that in many Mixtec languages the Irrealis mood is marked with a prefix which has two underlying forms: /k-/ and /ku-/. Their distribution is generally not predictable. The Realis mood prefix has one underlying form (except for an occasional irregular verb), which varies throughout the family. In Nuyóo it is /xi-/ and in Tezoatlán and Nuxáa it is /si-/ (where the s is phonetically [š] before front vowels and sometimes all non-low vowels).

In the following sections we discuss the mood morphology of the Mixtec verb as it is found in Nuxáa, Nuyóo, and Tezoatlán Mixtec. The discussion is organized around the three classes of verbs mentioned above.

3.1 Irregular consonant-initial roots

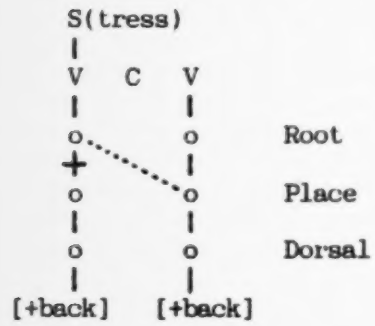
The verb roots in the first class which we discuss are all monosyllabic and begin with consonants. Some examples are given in (3). (The missing Tezoatlán forms are not cognate.) Other verbs which usually belong to this group include: -ta 'sing' and -nu^{a} 'run'.

(72)

	Nuyóo		Nuxáa		Tezoatlán	
	Irr.	Real.	Irr.	Real.	Irr.	Real.
'walk'	ka-ka	xi-ka	ka-ka	si-ka	ka-ka	si-ka
'throw'	ku-te?	xe-te?	ku-te	si-te	ku-te	si-te
'clear field'	ku-tu	xi-tu	ku-tu	si-tu		
'take care of'	ko-to	xi-to	ko-to	si-to		

We take the underlying form of the allomorphs ku-, ko- and ka- to be /ku-/. The surface forms are derived by the rule given in (73) which is a rule of complete vowel harmony (via the Root node) formulated in current feature notation (Clements 1985, Sagey 1986, Archangeli, to appear).

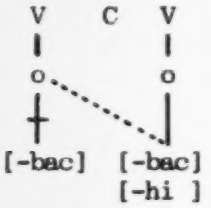
(73) Vowel Harmony (Nux, Nuy, Tez):



The reason for positing the prefix /ku-/ exclusively with these verbs is based on the observation that there are no inflected verbs of the shapes **ku-ka**, **ka-to**, or **ko-ku** superficially in these languages. The underlying form ku- is necessary for verbs like **ku-te** 'throw'. Rule (73) accounts for all of the data without positing any other underlying form.

In Nuyóo we see an additional vowel harmony rule operating in the Realis form, with i changing to e before a root with e. (This harmony rule may be collapsible with the previous rule, although we do not collapse the rules here.)

(74) Vowel Harmony (Nuy):



The phonological alternations are more radical with the verbs 'drink' and 'sleep', as shown in (75). No vowel remains constant and the two vowels in any verb form are identical. (Nuyóo also has an irregular allomorph of the realis prefix with the verb 'sleep': **ki-**.)

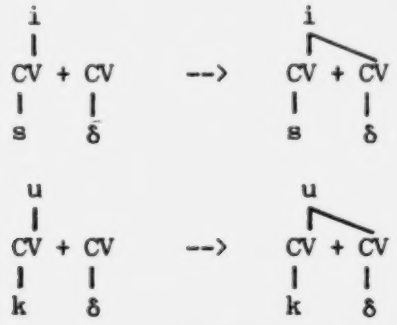
(75)

	Nuyóo		Nuxáa		Tezoatlán	
	Irr.	Real.	Irr.	Real.	Irr.	Real.
'drink'	ko?o	xi?i	ko?o	si?i	ko?o	si?i
'sleep'	kusu?	kisi?	kuðu	siði		

We propose that these verbs contain an underspecified vowel position in the root which is associated with the vowel of the prefix,

as shown in (8). This analysis assumes the segregation of vowels and consonants into different planes (see McCarthy 1986, to appear). The lowering of the u to o in the Irrealis form of 'drink' is not directly accounted for, however; another rule (not given here) is necessary.

(76)



The phonological alternations which these verbs illustrate are not shared by other consonant-initial roots. The irregularity of the members of this class of verbs appears to be related to the fact that they all have monosyllabic roots superficially. This is not true of the great majority of roots (verb or noun); Mixtec roots are typically bisyllabic superficially. If we assume that stress is usually assigned to the penultimate (or only) syllable of the root, then these roots are irregular in that they throw stress back leftward one more syllable, onto the prefix. The result is a stressed prefix vowel which does not undergo the deletion rules described for regular consonant-initial roots below.[18]

3.2 Vowel-initial roots

Vowel-initial verb roots are not as common as consonant-initial roots in Mixtec, but like the verbs of the previous class they display the mood morphology we are proposing. All of them are polysyllabic superficially. Most take the /ku-/ allomorph of the Irrealis morpheme (underlyingly); the exceptions are some a-initial verbs, some of which take /ku-/ (which surfaces as k^w) and some of which take /k-/. Some vowel-initial roots are given in (77), grouped by the initial vowel of the root. (Some gaps are due to the fact that the language in question does not have the cognate root. Others are due to the fact that the language does not have roots beginning with that vowel. In some cases, our data are simply incomplete.)

(77)		Realis	Irrealis	
	a.	i-initial		
	Nux	s-i?niⁿ	k-i?niⁿ	'tie'
	Nuy	x-i?ni?ⁿ	ku?-ni?ⁿ	'tie'
	Tez	s-i?o	k-i?o	'give'
	b.	e-initial		
	Nux	s-e?ne	k-e?ne	'harvest, cut'
	Nuy	x-e?ne?	k-e?ne?	'cut'
	Tez	(none)		
	c.	o-initial		
	Nux	s-oδo	k-oδo	'water'
	Nuy	y-oso?	k-oso?	'be on'
	Tez	s-o?niⁿ	k-o?niⁿ	'tie'
	d.	u-initial		
	Nux	(none)		
	Nuy	y-unu?ⁿ	k-unu?ⁿ	'love'
	Tez	(none)		
	e.	a-initial		
	Nux	s-eδɪ	k-aδɪ	'close'
	Nuy		k-asɬ	'close'
	Tez	s-aθi	k-aθi	'close'
	Nux	s-e?niⁿ	k-a?niⁿ	'kill'
	Nuy		k-a?niⁿ	'kill'
	Tez	s-a?niⁿ	k-a?niⁿ	'kill'
	Nux	s-esi	k-asi	'eat'
	Nuy	y-axi	k-axi	'eat'
	Tez	s-asi	k-asi	'eat'
	Nuy	x-aku?	kʷ-eku?	'laugh'
	Tez	s-aki	kʷ-aki	'laugh'
	Nux	s-e?eⁿ	kʷ-a?aⁿ	'go'
	Nux	s-e?e	kʷ-a?a	'give'
	Nuy	x-a?a	ku?-wa	'give'
	Nux	s-eniⁿ	k-aniⁿ	'lay'
	Nux	s-eδe	k-aδa	'do'

Nux	s-ee	k-aa	'rise'
Nuy	x-e?ne?^{n}	k-e?ne?^{n}	'cut'

The allomorph **y** of the Realis prefix appears only in Nuyóo (of these three languages) and only with verbs which begin with round vowels. We take it as a suppletive allomorph and do not discuss it further here.

In Nuyóo and Tezoatlán, as in most other Mixtec languages apparently, the vowel of the Realis prefix (/xi-/ (Nuy) or /si-/ (Tez)) simply deletes before a vowel-initial root. The rule, given in (78), also applies in Nuxáa, although it interacts with another rule that we discuss below. (Under assumptions of Strict Cyclicity, rule (78) applies only to derived vowel clusters.)

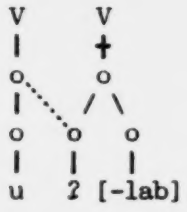
(78) Vowel Deletion (Nux, Nuy, Tez):

V --> Ø / ___ V

The **u** of the prefix /ku-/ does not delete by rule (78), however; underlying /ku-aki/ 'Irr-laugh' (Tez) surfaces as **k^{w}aki**. We assume that the prefix /ku-/ syllabifies as **k^{w}** before rule (78) applies in Nuxáa and Tezoatlan.

In Nuyóo the sequence /ku-V?.../ surfaces as **ku?** (see the verb 'give' and 'tie'); the vowel of the root simply deletes by a rule which applies before Vowel Deletion (78). This rule is given in (79), where ? stands for the feature [constricted glottis] in the laryngeal tier. This feature is retained although the root-initial vowel is lost.

(79) Root Vowel Deletion (Nuy):

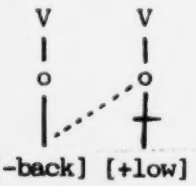


The surface forms of Irrealis forms of verbs beginning with u and o are predictable regardless of which underlying form is chosen (/k-/ or /ku-/). Since Mixtec has a general ban on two tautosyllabic labial segments, underlying sequences /ku-u.../ and /ku-o.../ could not surface as **k^{w}-u...** and **k^{w}-o...**; in both cases the Irrealis prefix surfaces simply as **k**.

Stems beginning with **a** in Nuxáa are e-initial in the Realis form. We assume here that there is spreading of the feature [-back] from the

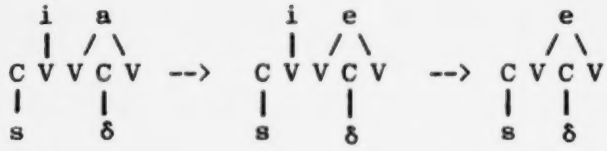
i to the a and that the prefixal vowel deletes by Vowel Deletion (78). This spreading rule is shown in (80).

(80) Raising (Nux):



When both syllables of a root contain the vowel a, as with the root /-aδa/ 'do', both vowels become e in the Irrealis. This alternation is readily accounted for without an extra rule if we assume the Obligatory Contour Principle and a multiplane analysis of consonants and vowels (McCarthy 1986, to appear). Under these assumptions, the features which constitute the vowel a appear only once in the root for 'do' since the vowels and consonants are segregated on different planes and contiguous identical elements are prohibited. The analysis is illustrated in (81).

(81)



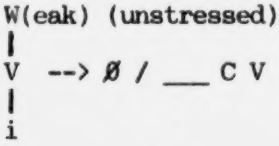
3.3 Regular consonant-initial roots

The most common verb roots in Mixtec belong to what we call the regular consonant-initial class. They do not throw the stress back, unlike the verbs in the irregular consonant-initial class, and these roots are superficially always polysyllabic. (See note 18.)

The most notable fact about the mood morphology with these verbs is that it often does not show up at all. One analysis would be to posit suppletive zero allomorphs of the mood prefixes for these verbs. Another analysis (more in line with what previous analysts have done) is to take the entire verb form as an indivisible form. Another alternative, which we adopt here, posits the same underlying forms as for the previous verbs and takes advantage of the phonological shape and prosodic structure of these words to account for the surface forms. Our analysis makes use of Stray Erasure, where unlicensed segments are deleted at the end of the phonological cycle. For example, it is predictable that the prefix /k-/ will not surface in Mixtec before consonant-initial roots since it cannot be syllabified. (Most Mixtec languages have a simple CV syllable as the maximal syllable, and none licenses consonant clusters such as kt or kn.) If we assume that in most languages an unstressed pretonic i is deleted, the deletion of the

fricative consonant of the Irrealis prefix is also explained by Stray Erasure. Stray Erasure is only circumvented if some other rule, such as an epenthesis rule, applies so as to create a structure in which the stray consonant will be licensed by the prosodic structure. Therefore we posit the following deletion rule for all three varieties of Mixtec:

(82) Deletion (Nux, Nuy, Tez):



This rule deletes only segmental material, not tones. In fact, the tonal evidence seems to support this 'abstract' analysis for regular consonant-initial roots. It appears to be the case that the Realis prefix is composed of the phonemic melody /si-/ (or its cognate) and a high tone (or its cognate -- our personal knowledge of Mixtec tone systems is very slight). The segmental material may delete under the conditions described above but the tonal material remains and is associated with the verb root, causing some widely attested tone sandhi in which high tone participates. (An alternative analysis for this tone is sketched out in the following section.) Some typical examples of regular consonant-initial roots are given below.

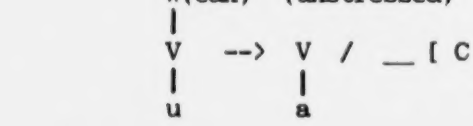
(83)	(Tez)	Realis	Irrealis	
	a.	kaniⁿ	kaniⁿ	'hit'
		θiko	θiko	'sell'
		noto	noto	'wake up'
		kawa?a	kawa?a	'do'
	b.	naka	ka-naka	'take care of'
		nati	ka-nati	'wait for'
		kuu	ka-kuu	'be'

(84)	(Nuy)	Realis	Irrealis	
	a.	sa?a	sa?a	'do, make'
		tutu?	tutu?	'whistle'
		ka?a	ka?a	'talk'
	b.	neeⁿ	ku-neeⁿ	'hold in arm'
		yatu?	ku-yatu?	'hurt'
		niso	ku-niso	'carry'

We believe that these roots indirectly display both allomorphs of the Irrealis prefix; underlying /k-/ surfaces as null (deleting by Stray Erasure), and underlying /ku-/ as ka- in Tezoatlán by rule (85). (This rule might alternatively be formalized as a delinking of the vowel

features which are specified underlyingly; the default feature specification would yield the surface vowel a.)

(85) Lowering (with unrounding) (Tez):



We do not take the vowel a as epenthetic for two reasons. First, in other Mixtec languages, (e.g. Ñumí), the u usually lowers to o. Therefore we know that a lowering rule is operative in some languages. Second, no epenthetic vowel occurs in the Realis form. Therefore we have no reason to believe that Stray Epenthesis (the epenthesis of a vowel before a stray consonant) is operative in Tezoatlán, unlike in Nuxáa (see below).

The regular consonant-initial verbs of Nuxáa divide into two groups. In the larger group, Stray Epenthesis is also not operative, as in the other two languages. Examples are given below.

(86) (Nux)	Realis	Irrealis	
	kaniⁿ	kaniⁿ	'hit'
	δiko	δiko	'sell'
	noto	noto	'wake up'

In the smaller group of verbs in Nuxaa, Stray Epenthesis is operative.[19] The inserted vowel is e after s, and a after k.[20]

(87) (Nux)	Realis	Irrealis	
	se-čii	ku-čii	'bathe'
	se-wa?a	ka-wa?a	'make'
	se-tone	ka-tone	'try'
	se-nodo	ka-nodo	'put'
	se-δiniⁿ	ka-δiniⁿ	'have supper'

3.4 Perfective and Imperfective aspect

In many Mixtec languages Perfective morphology consists of the (usually optional) morpheme niⁿ (or its cognate) and a tone (Mid in some languages, Low in Nuyóo) which is associated with the verb (first syllable in some languages) and which is not optional. The fact that the presence of the tone is not dependent on the presence of the morpheme niⁿ seems to indicate that the tone alternations in Perfective

forms are not due to a tone stranded by the deletion of segmental material.

The morpheme ni^{n} has sometimes been treated as a prefix and sometimes as a separate word, and it is certainly possible that it must be analyzed differently in different languages. We believe that it is probably best in most cases to take it as a phonologically dependent word, a clitic. In most languages it apparently cliticizes to the verb which follows it. In Nuyóo, however, it cliticizes to the preceding word since a pause may follow it but may not precede it.

We assume that Imperfective aspect is not marked directly. As mentioned in the previous section, a high tone typically occurs on Imperfective forms. We take this tone to be directly related to the high tone which occurs with the prefix /si-/ or its cognate and assume that Perfective aspect displaces it. An alternative analysis would take the High tone of the Imperfective form as a morpheme indicating Imperfective aspect and which is mutually exclusive with the tone indicating Perfective aspect. It is likely that only a very complete analysis of the tone sandhi patterns in Mixtec would allow us to choose between these alternatives.

Examples of Perfective and Imperfective aspect are given below. Surface tones are marked for Nuyóo and Tezoatlán, but not for Nuxáa. Unfortunately, analyses of the tone patterns in these languages are just beginning; therefore we are able only to point out that these languages follow the general patterns sketched above.

(88)

	Imperfective	Perfective		
(Nux)				
a.	s-i?ni^{n}	ni^{n} s-i?ni^{n}	'tie'	
	si-ka	ni^{n} si-ka	'walk'	
(Nuy)[21]				
b.	x-á?a	ǹ x-à?a	'give'	
	x-í?ní?^{n}	ǹ x-i?ni?^{n}	'tie'	
	xí-ka	ǹ xì-ka	'walk'	
(Tez)				
c.	s-a?ní^{n}	nì^{n} s-à?ni^{n}	'kill'	(Irr: k-a?ni^{n})
	sí-kà	nì^{n} sì-kà	'ask for'	(Irr: ka-kà)
	s-í?o	nì^{n} s-ì?o	'give'	(Irr: k-ì?o)
	káni^{n}	nì^{n} kani^{n}	'hit'	(Irr: kani^{n})

NOTES

[1]Mixtec languages are spoken in various regions in south central Mexico, chiefly in the state of Oaxaca. Together with Cuicatec and Trique, they form the Mixtecan family of the Otomanguean phylum (Longacre 1957, Campbell 1979). The various Mixtec languages are frequently referred to as 'dialects' of Mixtec; however, mutual intelligibility in different parts of the Mixtec region is sufficiently low to establish many of them as distinct languages (Egland 1972:25-37).

[2]This collaboration began in a workshop which we directed at the Summer Institute of Linguistics in Tucson in the spring of 1988. The people listed in (1) are conducting on-going research on these languages under the auspices of the Institute. This paper has also benefited from many helpful suggestions and comments by Barbara Hollenbach. Finally, we wish to acknowledge the contribution of those many native speakers who have generously and patiently shared their knowledge with us and our colleagues.

We have attempted to use a common orthography for all three languages, whose phonologies are for the most part very similar. However, the tonal systems of the three languages differ radically, and are not sufficiently analyzed at this point to make definitive statements. Indeed, the system in Nuxáa is too complex to even notate tone at this point, and we have provisionally adopted a transcription system that assumes three tones in Nuyóo and Tezoatlán (´ high, ` low, mid unmarked), although there is a possibility of a fourth tone in both systems. No conclusions should be drawn from our transcription of tone, beyond those that we indicate in the text.

Through the Mixtec family, all morphemes are classified in the lexicon either as nasal or oral (Marlett, in preparation); we have written nasal morphemes with a superscript final 'n' (^{n}). Nasalization of a morpheme is realized phonetically by spreading leftward through sonorants (with a few exceptions) from the end of the morpheme. Besides nasalizing vowels, this spreading also selects the nasal allophones of three consonants.

	Oral words	Nasal words
/n/	n^{d}	n
/y/	y ~ ž	n^{y}
/w/	β ~ v	m

The only exception to this is in Nuyóo, where [n] occurs in both nasal and oral words as the only allophone of /n/; there is no [n^{d}]. Low level phonetic rules sometimes obscure the effects of nasal spreading, either nasalizing or removing nasalization from vowels adjacent to nasal consonants. For more details, see Williams and Williams 1988 (Tezoatlán) and Hugghins and McKendry 1988 (Nuxáa).

Clitics abound in these languages. For the most part we have written them as separate words, sometimes adding an equal sign (=) to indicate phonological dependency. This is especially true for pronominal clitics, which occupy syntactic positions such as subject, object, and possessor, but are phonologically dependent on the items to their left. The clitic pronouns may, under some circumstances, be replaced with various types of free pronouns, some indicating emphasis, others simply providing more semantic information (such as number) not available in the clitic pronouns. Clitic pronouns are glossed with a number indicating person, followed by one or more of the following abbreviations:

s	singular (non-honorific)
p	plural (non-honorific)
1pi	1st plural inclusive
1pe	1st plural exclusive
r	respect (singular or plural)
m	masculine (adult)
f	feminine (adult)
n	neuter
c	child
a	animal
i	inanimate
Indef	indefinite

Free pronouns are glossed either by the above abbreviations, or by the closest English equivalent.

The following additional abbreviations are the most important ones used in glosses. (The meanings of others are obvious from context.)

Hab	Habitual
Imp	Imperative
Impf	Imperfective
Irr	Irrealis
Neg	Negative
Opt	Optative
Perf	Perfective
Prog	Progressive

[3]Throughout, we use capitalized words (e.g. 'Perfective', 'Aspect') to refer to language-particular grammatical categories, i.e. convenient labels attached to specific elements in Mixtec. Words entirely in lower case (e.g. 'perfective', 'aspect') are used to refer to categories defined independently of individual languages. Our goal is to characterize the language-specific categories in terms of the universal ones.

[4]Throughout, we use 'situation' as a cover term for events, states, and processes.

[5]This characterization of the perfective/imperfective distinction is based on Comrie (1976:16-24). There, he argues for this characterization as opposed to various others, such as punctiliar vs. durative, short vs. long (in absolute terms), or completed vs. not (yet) complete. Indeed, Nuxáa has explicit morphology which indicates a situation as not yet complete, which is independent of the perfective/imperfective distinction. For an alternative characterization of perfectivity in terms of 'closure', see Chung and Timberlake (1985:217,219).

[6]We have found no uses of the Perfective with present time reference. However, we attribute this to the fact that (virtually) all situations occupy some time span with nonzero length, while the present moment is apparently treated linguistically as having zero length. It is extremely unlikely that any presently occurring situation would be represented as a complete whole, or viewed 'from the outside in'. The prominence of that portion of the situation which coincides with the present moment is simply too great for it to be ignored, as it would with a Perfective form. Even if there are verbs which are inherently momentary, such as verbs of arrival (cf. Macaulay 1985), it would be difficult to distinguish present reference of a perfective form from immediate past reference, e.g. 'He arrives' vs. 'He just arrived'.

[7]Similarly, as Barbara Hollenbach has pointed out, 'Imperfective' should not be confused with the 'Imperfect' in languages such as Spanish, which in our terms is a *past* imperfective.

[8]Inasmuch as these are motion verbs, the normal way to achieve habituality is by iteration, so that either label is appropriate. Further, since these verbs are dynamic, the Progressive could equally be called 'Continuous'. (On the importance of dynamicity to the distinction between progressive and continuous, see Comrie (1976:32-35) and Chung and Timberlake (1985:218-9).) We avoid the term 'Continuous', because of the frequent use of 'Continuative' in Mixtec studies to designate the Imperfective.

[9]Kent Wistrand apparently also noticed it about this time, but we don't know who discovered it first.

[10]In Tezoatlán, Perfective forms are sometimes used for one-way motion. That is, motion verbs in this language seem to be developing senses in which return to the point of origin is not important.

[11]This provides a further reason for not confusing the term 'Continuative' (= Imperfective) with 'continuous', inasmuch as 'continuous' means 'non-habitual'. That is, the Continuative of motion verbs is specifically not continuous.

[12]It may seem odd to include future time as one instance of a category of mood. However, the usage is justified in that many languages do not distinguish between irrealis mood and future tense. Even in English, future tense has an extended sense to refer to present situations which are very likely but not known to be actual.

John will (most likely) be married by now.

If 'mood' is defined as morphology which expresses the actuality of a situation, then future can plausibly be seen as one type of nonactuality. Languages, then, differ in what is classified as nonactual for the purposes of a particular modal system, and thus terms like 'realis' and 'irrealis' will always require definition with respect to a particular language. (Cf. the discussion in Chung and Timberlake 1985:243ff.)

[13]We have not been able to discover all the factors that determine whether the subject will be overt or absent in particular sentences. When absent, it is always understood to be a second person pronoun; i.e., only second person pronouns can be omitted. However, the exact details vary considerably in the three languages, due in part to the differences in their pronominal systems.

[14]The relative future sense of the Irrealis also comes out in some discourses, where the reference point from which the actuality of a proposition is judged is apparently in some other sentence. For example, the following is from a text describing former marriage customs.

kenta =raa kukʷéntu =raa
Irr/arrive 3p Irr/chat 3p
They would arrive and chat

xiíⁿ tátá? ya?aⁿ. (Nuy)
with father woman
with the father and mother.

Such uses are not yet well understood.

[15]By presupposition, we mean that portion of the meaning of the whole sentence which is stable under negation and yes-no questioning. It is material that is assumed to be known and accepted as true by both speaker and hearer.

[16]In Nuxáa, the use of realis forms for presupposed situations can be overridden by including the optative particle **naᵐ**. In Nuxáa, this particle is frequently used in contexts where other Mixtec languages would not use it. It always seems to express the desirability of the situation, from the point of view of the speaker, and always takes the Irrealis.

ki naⁿ kuu =nuⁿ ha kučii =nuⁿ,
when Opt finish 3f Comp Irr/bathe 3f
When she finishes bathing,

δate ki?iⁿ =niⁿ ha kučii =niⁿ. (Nux)
then Irr/go 1 Comp Irr/bathe 1
then I can go (to) bathe.

[17]Certain clauses in Nuyóo seem to have both a presupposed temporal ('when') interpretation and a hypothetical ('if') one. At any rate, we have not been able to determine any way to distinguish the two interpretations in conversations with our Nuyóo consultant. In both interpretations, Irrealis forms are used in the subordinate clause.

ré kikawa =nì? yakʷáá? wita
if Irr/lie.down 1 night today
If/when I lie down tonight,

(soo) kusu wà?a =nì?. (Nuy)
but Irr/sleep good 1
I will sleep well.

It may be that 'when' clauses simply do not exist as such, and hypothetical 'if' clauses have expanded their meaning slightly to fill the void. Whatever the explanation, this represents a possible exception to the analysis given in the text. Note, however, that in Nuyóo temporal clauses beginning with **ni** 'while', realis forms are used, as in (57).

[18]Regular consonant-initial roots may also be monosyllabic underlyingly, in our view, but they retain the stress and undergo a lengthening rule. This rule changes underlying CV to CVV and underlying CV? to CV?V.

[19]This division of the lexicon is reminiscent of that found in Zapotec languages (Marlett and Pickett 1987).

[20]An alternative which is worthy of consideration takes the **se** as an allomorph of /si-/ by a lowering rule rather than by epenthesis. Another alternative which might be considered would posit roots beginning with **a** for verbs like 'make', /-awa?a/. Raising (80) and vowel deletion (78) would account for most of the surface allomorphy. The main problem with this alternative is that it does not account for the **e** in the Realis form of the verb 'bathe'.

[21]The segment **x** is often deleted phonetically when it is preceded by the Perfective morpheme **n** in Nuyóo.

REFERENCES

Alexander, Ruth Mary. 1980. Gramática mixteca de Atatláhuca. Gramáticas de lenguas indígenas de México 2. Instituto Lingüístico de Verano, México.

Archangeli, Diana. To appear. Constraining phonological rules: Implications for feature organization. In Bruce Hayes, ed. Segment structure.

Bradley, C. Henry. 1970. A linguistic sketch of Jicaltepec Mixtec. SIL Publications 25. SIL, University of Oklahoma, Norman.

Bradley, C. Henry, and Barbara E. Hollenbach, eds. To appear. Studies in the syntax of Mixtecan languages. (3 vols.) Summer Institute of Linguistics publications in linguistics, SIL and the University of Texas, Arlington.

Bybee, Joan. 1985. Morphology: A study of the relation between meaning and form. John Benjamins, Amsterdam.

Campbell, Lyle. 1979. Middle American languages. In Campbell, Lyle, and Marianne Mithun, eds., The languages of native America: historical and comparative assessment. University of Texas Press, Austin. Pp. 902-1000.

Chung, Sandra, and Alan Timberlake. 1985. Tense, aspect, and mood. In Shopen, Timothy, ed. Language typology and syntactic description, vol. III, Grammatical categories and the lexicon. Cambridge University Press, Cambridge.

Clements, George. 1985. The geometry of phonological features. Phonology yearbook 2: 223-250.

Comrie, Bernard. 1976. Aspect. Cambridge University Press, Cambridge.

Crystal, David. 1985. A dictionary of linguistics and phonetics. 2nd edition. Basil Blackwell, New York.

Daly, John P. 1973. A generative syntax of Peñoles Mixtec. SIL Publications 42. SIL, University of Oklahoma, Norman, OK.

Dyk, Anne, and Betty Stoudt. 1965. Vocabulario mixteco de San Miguel el Grande. Vocabularios Indígenas 'Mariano Silva y Aceves' 12. Instituto Lingüístico de Verano, México.

Egland, Steven T. 1978. La inteligibilidad interdialectal en México: Resultados de algunos sondeos. Instituto Lingüístico de Verano, México.

Gittlen, Laura and Stephen A. Marlett. 1985. Ñumi Mixtec syllable structure and morphology. 1985 Workpapers of the Summer Institute of Linguistics, University of North Dakota Session, 175-194.

Hugghins, Susan, and Inga McKendry. 1988. Nuxáa Mixtec Phonology. Paper presented at the Summer Institute of Linguistics, Tucson AZ, April 23, 1988.

Josserand, Judy Kathryn. 1983. Mixtec dialect history. Dissertation, Tulane University.

Kuiper, Albertha and William R. Merrifield. 1975. Diuxi Mixtec verbs of motion and arrival. IJAL 41:32-45.

Longacre, Robert E. 1957. Proto-Mixtecan. IJAL 23:4 supplement.

Macaulay, Monica. 1985. The semantics of 'come', 'go', and 'arrive' in Otomanguean languages. Kansas working papers in linguistics, 10(2):56-84.

Macaulay, Monica. 1987. Subordination and the subjunctive in Chalcatongo Mixtec. Conference on American Indian languages, November 22, 1987.

Marlett, Stephen A. In preparation. Nasalization in Mixtec.

Marlett, Stephen A. and Velma B. Pickett. 1987. The syllable structure and aspect morphology of Isthmus Zapotec. IJAL 53:398-422.

McCarthy, John. 1986. OCP effects: Gemination and antigemination. Linguistic Inquiry 12:373-418.

McCarthy, John. In preparation. Planes and linear ordering.

North, Joanne and Jäna Shields. 1976. Gramática popular del mixteco de Silacayoapan. Instituto Lingüístico de Verano, México.

Pensinger, Brenda J. 1974. Diccionario mixteco-español español-mixteco. Vocabularios Indígenas 'Mariano Silva y Aceves' 18. Instituto Lingüístico de Verano, México.

Sagey, Elizabeth. 1986. The representations of features and relations in nonlinear phonology. Dissertation, MIT, Cambridge.

Speck, Charles J. and Velma B. Pickett. 1976. Some properties of the Texmelucan Zapotec verbs 'go', 'come', and 'arrive'. IJAL 42:58-64.

Stark Campbell, Sara, Andrea Johnson Peterson, and Filiberto Lorenzo Cruz. 1986. Diccionario mixteco de San Juan Colorado. Vocabularios Indígenas 'Mariano Silva y Aceves' 29. Instituto Lingüístico de Verano, México.

Williams, John L., and Judith Williams. 1988. Tezoatlán Mixtec Phonology. Paper presented at the Summer Institute of Linguistics, Tucson AZ, April 23, 1988.

NONSYLLABIC VOCOIDS*

John M. Clifton

1 Introduction
2 Factors resulting in nonsyllabic vocoids
  2.1 Unmarked consonant position
  2.2 Relative vowel height
  2.3 Labialization and palatalization
3 Factors blocking nonsyllabification of vocoids
  3.1 Stress
  3.2 Morpheme breaks
  3.3 Long vowels
4 Conclusion
Notes
References

1 Introduction

It has long been recognized that syllabic [i u] and nonsyllabic [y w] are related. In discussing Saussure's Mémoire sur le système primitif des voyelles dans les langues indo-européennes, Anderson (1985) notes that:

> "The most famous aspect of this system was the theory of 'coefficients sonantiques'; these were the elements, like the liquids and nasals and the high vowels/semivowels (i/y, u/w), which could be realized either as vowels (when surrounded by consonants) or as consonants (when preceded by a vowel)." (p. 20)

Anderson also discusses Saussure's idea that a universal inventory of phonetic segments, called 'phonetic species', have various manifestations, called 'phonemes', depending on their position within the syllable.

> "Specific phonemes are the positional realizations of phonetic species, where the variations among them are due primarily to general phonetic, rather than language-particular, principles. Fundamentally, it is syllabic organization that is being idealized away from here... For example, one of the recurring examples of different phonemes belonging to the same species is Saussure's description of (prevocalic, onglide) [y], (vowel) [i], and (postvocalic offglide, or second element of a diphthong) [i̯] as members of the same species [I]." (Anderson 1985:39)

Working within a framework in which segments are interpreted on a language specific basis, Pike (1947:129) arrives at similar conclusions.

> "The sound types just listed [high vocoids, retroflexed central vocoids, voiceless vocoids, lenis voiced velar fricatives - JMC] will function as consonants or vowels according to the place in which they occur in basic phonological and grammatical units of the particular language under consideration, and according to the structural pressure exerted upon them in these positions by structural analogies."

Chomsky and Halle (1968:408), on the other hand, attempt to characterize the universally preferred consonant positions.

> "...in word initial position the conventions select true consonants as the unmarked segments, vowels as the fully marked segments, and glides and liquids as intermediate in complexity. After a consonant, vowels are fully unmarked and glides fully marked. After a vowel, true consonants are fully unmarked and vowels fully marked. We might easily extend these rules to distinguish additional cases."

Working within a nonlinear theory of syllable structure, Mohanan (1986:29-31) proposes an analysis of syllabic consonants and semivowels in English following Kiparsky's (1982) suggestion to return to the position in Chomsky and Halle (1968) in which all redundant information is eliminated from underlying representations. Mohanan's analysis of semivowels is that:

> "the nonsyllabic /i/ and /u/ in words like <u>yam</u> and <u>wet</u> are marked in the underlying representations as nonheads [i.e. [-syll][1] - JMC], and all other segments are unspecified for syllabicity..." (1986:30)

The nonspecified segments that are vocalic ([+son, -cons]) are then marked as heads, that is, [+syll]. In other words, Mohanan claims (at least for English) that the unmarked state for high vocoids is [+syll].

In this paper, I will examine a number of New Guinean languages in an attempt to show that this analysis is too simplistic. When determining the markedness of a given segment, it is necessary to examine both paradigmatic and syntagmatic relationships. Paradigmatically, it is undoubtedly true that a nonconsonantal sonorant (i.e. vocoid) unmarked for syllabicity is [+syll]. Syntagmatically, however, there are a number of environments in which a vocoid unmarked for syllabicity is [-syll]. In section 2, I discuss a number of environments in which the unmarked status of a vocoid is [-syll]. Then in section 3, I outline factors which interact with these environments, resulting in syllabic segments where nonsyllabic vocoids are more natural.

2 Factors resulting in nonsyllabic vocoids

2.1 Unmarked consonant position

It is commonly asserted that certain positions within a phonological string are generally filled by consonants, while others are generally filled by vowels. If vowels are [+syll], while glides are [-syll], Chomsky and Halle's observations referred to in the Introduction involve the claims that the unmarked feature for high vocoids is [+syll] after a consonant, but [-syll] in word initial position and after a vowel. These observations capture the distinctions evident in the following data from Kope (Clifton 1987a:4).

(1) [idi$\underline{\text{i}}$so] 'to drink'

(2) [m$\underline{\text{e}}$$a] 'good'

(3) [$^{\text{i}}$amagauri]^{2} 'to jump'

(4) [ma$$^{\text{u}}$a] 'sibling'

The underlined vocoids in (1-2) are syllabic since they follow consonants. In (3) the [$^{\text{i}}$] is nonsyllabic since it is word initial, and in (4) the [$^{\text{u}}$] is nonsyllabic since it follows a vowel.

The phenomenon illustrated in the above discussion can be accounted for by the following processes.[3]

Word Initial Nonsyllabification (WIN)[4]

A word initial high vowel is [-syll] before another vowel.

Post Vowel Nonsyllabification (PVN)

An intervocalic high vowel is [-syll].

The term 'nonsyllabification' indicates that an unmarked vocoid does not receive the default specification [+syll]. It is important to note that neither WIN nor PVN refer to syllable structure. Mohanan (1986), following Clements and Keyser (1983), argues that syllable structure is predictable given that each segment is marked for the feature [syll]. Obviously, the value for the feature [syll] cannot be determined for a given segment in relation to the syllable structure if this syllable structure is dependent on the value for [±syll] for each segment.

It is also interesting to re-examine the English words wet and yam, as opposed to Iago. As mentioned in the Introduction, Mohanan (1986) claims the initial glide in wet and yam must be marked in the lexicon as [-syll] on the assumption that the unmarked state for vocoids is always [+syll]. The syllabic [i] in Iago, on the other hand, is

unmarked for Mohanan. These claims seem to be counterintuitive - Iago, not yam, seems to be the aberrant form in English. These intuitions are predicted by WIN since it claims word initial vocoids are naturally nonsyllabic.

One other situation not accounted for by WIN or PVN is found in languages in which closed syllables are allowed. Because there are no closed syllables in Kope, the unmarked value for a vocoid after a consonant is [+syll]. The following data from Mianmin (Smith and Weston 1974) illustrates the situation is different in a language which allows closed syllables.[5]

(5) /was/ 'drum'

(6) /awitnin/ 'star'

(7) /ikamwal/ 'married couple'

(8) /amun/ 'stomach'

The nonsyllabic nature of the /w/ in (5-6) is predictable from WIN and PVN, respectively. The nonsyllabic /w/ in (7), on the other hand, cannot be accounted for by either process. Since, however, closed syllables are allowed in Mianmin both in word final position, as in (5-8), and in word medial position, as in (6), the nonsyllabic nature of the /w/ in (7) is to be expected. The /u/ in (8), on the other hand, must be syllabic since there is no vocoid following it to become the peak of the syllable. This situation is accounted for by the following process.

Post Consonant Nonsyllabification (PCN)

A high vocoid is [-syll] when following a consonant and preceding a vowel.

Just as WIN and PVN do not refer to syllable structure, PCN does not refer to syllable structure. PCN will only be found, however, in languages in which closed syllables are permissible. This is because the nonsyllabic /w/ in a form like (7) means that the preceding consonant must either close the preceding syllable, or be part of a syllable initial consonant cluster. Since in Kope neither syllable type is possible, a form like (7) would be impossible. In a language allowing closed syllables, on the other hand, the application of PCN is not only possible, but expected.[6]

In summary, then, the unmarked value for vocoids is nonsyllabic before another vocoid in the following positions: word initial; after a vocoid; and, in languages with closed syllables, after a consonant.

2.2 Relative vowel height

McElhanon (1967) shows that in Selepet syllabification of vocoid sequences is determined by the relative articulatory position of the members of the sequence. Selepet has six vowels as indicated in (9).

(9)
i	u
e	o
a	ɔ

These six vowels can pattern in sequences with few restrictions. Representative examples involving /a/ and /e/ are shown in (10-16).

(10) /ie ei/ a. [si$ep si$ep] 'kind of bird'
b. [mem kepeˡ] 'hold it and twist it!'

(11) /ia ai/ a. [si$ap] 'it burned'
b. [kaˡde] 'kind of weed'

(12) /ue eu/ a. [lu$e lu$e ˡap] 'it howled'
b. [geᵘnaᵒ] man's name

(13) /ua au/ a. [huhu$aksap] 'he smashed himself'
b. [baᵘ] 'pig'

(14) /ea ae/ a. [he$ak] 'breath'
b. [gaᵉk] 'let him come!'

(15) /oa ao/ a. [go$aksɔm] 'bending over'
b. [kaᵒk] 'white'

(16) /eɔ ɔe/ a. [gɔre$ɔk] 'let her sew!'
b. [sɔhɔᵉk] 'let him tie it!'

In (10a-16a) the sequence syllabifies as two syllables, in line with the claim that vocoids unmarked for syllabicity should be syllabic. In (10b-16b), however, the sequence consistently syllabifies as one syllable. The difference between the (a) and (b) forms is that the first vowel is higher than the second in (a) but lower than the second in (b).

One way to account for this data is to claim that the second vocoid in sequences like /au ao/ is nonsyllabic. This would involve a process that was essentially the opposite from those proposed in section 2.1 in which the first vocoid in the sequence was nonsyllabic. A second way to account for this data is by the following process of Low-High Syllabification.

Low-High Syllabification (LHS)

The sequence V_1V_2 syllabifies as one syllable if V_1 is lower than V_2.

This process does not claim that V_1 is syllabic while V_2 is nonsyllabic. Instead, it claims that the sequence V_1V_2 makes up a single syllabic segment, similar to [ᵐb] being a single prenasalized stop. I will assume the second analysis since there are some basic differences between the situation in Selepet and the situations presented in section 2.1.

First, in the examples in section 2.1, the nonsyllabic vocoid always occurred in positions which could also be filled by a consonantal segment. This is not the case here. The nonsyllabic [ᵘ] in [baᵘ] 'pig' presents no problem since there are syllable final consonants in Selepet. The nonsyllabic [o] in [kaᵒk] 'white', however, does present problems since there are no instances of consonant clusters in syllable final position. A similar situation arises in Kope (Clifton 1987a). Although there are no syllables closed by consonants in Kope, there are sequences similar to those seen in Selepet as illustrated in (17-18).

(17) [meⁱ$veⁱ] 'pineapple'

(18) [mi$haᵉ] 'rain'

Pike (1947:130) notes a similar problem in English.

> "...a difficulty occurs in English, with nonsyllabic, postsyllabic [i] or [ɪ] in [ai], [au], [ɔi]. It is difficult to determine whether the second element of one of these sequences is a consonant or a vowel. One might affirm that [rai] 'rye' is analogous to [rat] 'rot', and so the [-i] is phonemically /-y/; but one also finds [rait] 'write' in which one sees that the [-i] does not actually parallel the [-t], after all, since one does not find *[ratt]."

Later, he suggests this may be an instance of a 'close-knit nucleus':

> "A syllable may have a CLOSE-KNIT nucleus composed of two vocoids... which acts in distribution (in the syllable) like a single simple nuclear phoneme." (p. 148, emphasis in original)

Whether these sequences are analyzed as VS or as VV which act as a single syllable nucleus, a structure must be set up which has no independent motivation.[7] For the remainder of this section, I will refer to the phenomena outlined in section 2.1 as an SV pattern and the phenomena outlined in this section as a VV pattern.

A second characteristic differentiating the SV pattern from the VV pattern relates to the merger of phonological contrasts. In Selepet, the basic contrast between /e/ and /i/, and between /o/ and /u/ was maintained even in the VV patterns, as seen in (14b) vs. (11b) and in (15b) vs. (13b). Similar contrasts are maintained in Kope (Clifton, in progress) and Mengen (Rath 1981).

In none of these languages, however, are similar contrasts maintained in SV patterns. Instead, the S is always high. Moreover, Lichtenberk (1980) claims that in Manam /o/ raises to [w] in the environment for WIN, as shown in (19), or PVN, as shown in (19-20).

(19) /oaoa=oa/ [wawa=wa] 'white'

(20) /eoa/ [ewa] 'fire'

Turner (1986:21) makes the following observations about Manam.

> "The 'front' vowels, i and e change to sound like an English 'y' and the back vowels, u and o change to sound like an English w. Of course, Manams can usually tell the difference between i and e when they sound like 'y' and between u and o when they sound like 'w'..."

Similar claims have been made for Mengen by Rath (1981). Contrasts between /u/ and /o/ are merged when they are nonsyllabic due to WIN or Lab (to be discussed in section 2.3), but not when they are nonsyllabic due to LHS.[8]

Finally, D. James (personal communication) has indicated that in the SV pattern, the S is definitely shorter in duration than the V, which is perceptually clearly the core of the syllable. In the VV pattern, on the other hand, neither V is clearly shorter than the other, and neither is clearly the core of the syllable.[9] It seems, then, that the phenomena accounted for by LHS is different from the phenomena accounted for by WIN, PVN, and PCN.

LHS is evident in a wide variety of languages. In Indonesian, while the sequences /ia ua/ represent two syllables, the sequences /ai au/ represent one syllable. Gasaway and Sims (1977:24) state that in Girawa:

> "...those clusters with a high vowel /i,u/ as second member have a single beat of prominence, while those with a high vowel as first member have two beats of prominence."

Similarly, in Numanggang (Hynum 1987:4-5) /ie ia io ue ua uo/ occur with two mora of timing, while /ei ai oi eu au ou/ occur with "slightly more than one mora of timing."

LHS applies in a large number of New Guinean languages, although there also seem to be languages like Buin (Griffin 1986) in which it does not apply. In these languages all vowels in a sequence are separate morae.[10]

LHS and PVN can both apply in forms such as (4) from above, repeated here as (21).

(21) [ma$ᵘa] 'sibling'

According to PVN the [ᵘ] should be nonsyllabic because it is intervocalic. This results in the [ᵘ] being syllable initial as shown in (21). At the same time, LHS predicts the sequence [aᵘ] should be a single syllable nucleus. Clifton (1987b) claims that both processes apply, resulting in an ambisyllabic [ᵘ]. That is, the [ᵘ] is part of the syllable nucleus of the first syllable, and simultaneously the onset of second syllable.

R. Steinbring (personal communication) provides evidence from the Yale (or Kosarek) language which indicates that in cases like (21) the application of PVN bleeds LHS. The phonetic sequences [eⁱ aᵉ aᵒ oᵘ] are found as single syllable nuclei in Yale. The sequences [aⁱ aᵘ], however, are not found in any form. As shown in (22), when a sequence of /a+u/ arises across morpheme boundaries it becomes [oᵘ].

(22) /ia+ukap/ [ioᵘ$kap] 'we will come'

At the same time, sequences of [a$ᵘV] arise as shown in (23).

(23) [a$ᵘe] 'garden'

Both PVN and LHS generally apply in Yale. If LHS applied in (23), however, the initial [a] should be raised to [o]. Since it is not, PVN must first apply making the [ᵘ] nonsyllabic and bleeding LHS.[11] It is an empirical question as to whether in other languages both processes apply, resulting in ambisyllabic high vocoids. I have no evidence of such an application being necessary in any language, however.

A process related to LHS may apply to vowels of the same height. McElhanon gives the following examples in which both vowels are the same height.

(24) /iu ui/ a. [pi$u pi$u] 'kind of bird'
b. [puⁱ] 'chicken'

(25) /eo oe/ a. [gɔre$on] 'you sewed it'
b. [toᵉn] 'in the water'

Once again, in (24a-25a) each vowel in the sequence represents a separate syllable, while in (24b-25b) the two vowels syllabify together. In this case, the first vowel is front in the (a) forms but back in the (b) forms. This data can be accounted for by the following process of Back-Front Syllabification.

Back-Front Syllabification (BFS)

The sequence V_1V_2 syllabifies as one syllable if the two vocoids are the same height and V_1 is articulatorily farther back than V_2.

Variation can be seen between languages with regard to V_1V_2 sequences of equal height. In Girawa (Gasaway and Sims 1977:24) all V_1V_2 sequences syllabify as single syllables regardless of the backness of the two. The syllabification in forms like [tiᵘ] 'pandanas type' and [siᵘr] 'mosquito' provides motivation for the following process of Front-Back Syllabification.

Front-Back Syllabification (FBS)

The sequence V_1V_2 syllabifies as one syllable if the vocoids are the same height and V_1 is farther front than V_2.

Both BFS and FBS apply in Girawa, while just BFS applies in Selepet. Neither applies in Numanggang (Hynum 1987:5) where /iu/ and /ui/ both syllabify as two syllables as shown in [ki$ut] 'join' and [dobu$i] 'cut a trail'. There do not seem to be any languages in which FBS but not BFS applies, however. That is, there do not seem to be languages in which a sequence of front vowel followed by back vowel is generally syllabified as a single syllable, while the reverse sequence is syllabified as two syllables.

The sequences which routinely syllabify as single syllables sometimes coalesce into a single phone. In Kope (Clifton, in progress), the final word in a noun phrase takes /-i/ suffix. Thus, [geema] 'large' becomes [geemai] when phrase final. The word [pɛ] 'canoe', however, becomes [pe], not *[pɛⁱ]. In general, /e/ is realized as [ɛ], while /ei/ is realized as [e].

Rath (1981) states that in Mengen,

> "In some words the phone [o] is heard alternating with the sequence [au] in rapid speech.... This phone [o] will then be interpreted as an optional portmanteau phone of the sequence /a/ + /u/."

Rath (personal communication) also indicates that the phones [æ] and [ɔ] come from /ae/ and /ao/, respectively. This analysis is based on several factors. First, while vocoid clusters are common, there are no clusters of /ae/ or /ao/ within morphemes. Second, [æ] and [ɔ] do not cluster with other vowels. Third, there is morphophonemic evidence from reduplication, as shown in (26-28).

(26) a. [laᵘ$na] 'its leaf'
b. [la$laᵘ$na] 'its leaves'

(27) a. [kⁱu] 'to call'
b. [ki$kⁱu] 'calling'

(28) a. [tɔ] 'to walk'
b. [ta$tɔ] 'walking'

Reduplication is used for pluralization and continuous tense in Mengen. As shown in (26-27), only the first part of a sequence of vocoids is reduplicated.[12] The fact that [ɔ] in (28) is reduplicated as [a] indicates it is a diphthong beginning with /a/.

To summarize, sequences of V_1V_2 generally syllabify as one syllable when V_1 is lower than V_2, or, if they are the same height, when V_1 is farther back than V_2. Sequences that syllabify as one syllable may further coalesce to form a single phonetic segment.

2.3 Labialization and palatalization

In many languages /u/ and /i/ coalesce with a preceding consonant when followed by another vocoid, resulting in labialized and palatalized consonants, respectively. In a sense, these high vocoids are naturally nonsyllabic in these environments. In Imonda (Seiler 1985:15) all nonfinal labial and velar consonants are labialized when adjacent to back vowels. Before front vowels, however, there are contrasts between labialized and nonlabialized labial and velar consonants as seen in (29-30).

(29) a. [pete] 'coffee'
b. [pᵘeta] 'silent'

(30) a. [xe] 'cut'
b. [xᵘef] EMPHATIC

Alveolar consonants are never labialized, but they do occur before sequences of /uV/ as shown in (31-32).

(31) [sue] 'fire'

(32) [tui] 'butterfly'

Labial and velar consonants, on the other hand, never occur before sequences of /uV/. From this, we can conclude that labialized consonants arise from sequences of /CuV/ where the consonant is labial or velar. The phonemic forms of (29b) and (30b) are /pueta/ and /xuef/, respectively.[13]

In Mai Brat, Brown (1987) outlines a similar process of palatalization affecting /i/. Brown describes the process as follows.

> "...in the sequence C V1 V2 where C is either a noncontinuant consonantal or a strident /b, t, k, m, n, s/ and V1 is /i/ and unstressed, C V1 becomes C."

Examples of this are given in (33-34).

(33) /ni6/ [ñ6] 'you'

(34) /sióx/ [sʲóx] 'fish'

In Mengen (Rath 1981), all sequences of V_1V_2 are syllabified as one syllable as shown in (35-39).

(35) [ba$baʲ] 'to think'

(36) [gʲaᵘ] 'my name'

(37) [kᵒaʲ] 'this'

(38) [mᵘe] 'snake'

(39) [sᵒa$re] 'spirit'

LHS can account for the syllabification of [aʲ aᵘ] in (35-37). It cannot account for the syllabification of [ʲa ᵒa ᵘe] in (36-39), however. This syllabification is accounted for by Rath in terms of labialization and palatalization of the preceding consonant.

The labialization and palatalization illustrated above are accounted for by the following processes.

Labialization (Lab)

The vocoid /u/ (and possibly /o/) coalesces with a preceding consonant when followed by another vocoid, resulting in a labialized consonant.

Palatalization (Pal)

The vocoid /i/ (and possibly /e/) coalesces with a preceding consonant when followed by another vocoid, resulting in a palatalized consonant.

The actual class of consonants referred to by Lab and Pal will be defined on a language specific basis. They are formulated as two separate processes since they seem to occur independently in languages. Imonda exhibits only Lab, Mai Brat exhibits only Pal, and Mengen exhibits both.

3 Factors blocking nonsyllabification of vocoids

There are at least two specific factors which can block the nonsyllabification of vocoids: stress and morpheme breaks. In addition, the presence of long vowels in a language can result in apparent violations of the processes outlined above, so this will also be outlined in a separate section.

3.1 Stress

A stressed vocoid will generally not be nonsyllabic even if in an environment which would normally result in nonsyllabification. For example, it was mentioned in section 2.3 that palatalized consonants were created in Mai Brat only if the /i/ was unstressed. Stress is contrastive in Mai Brat, as shown by (40-41) from Brown (1987).

(40) a. [násom] 'you carry'
b. [nasóm] 'your name is'

(41) a. [máru] 'she cuts'
b. [marú] 'lake'

The Mai Brat forms with palatalized consonants given above as (33-34) are repeated here as (42a-43a), along with minimal pairs with stressed /i/ and no palatalization.

(42) a. /nió/ [ñó] 'you'
b. /nío/ [nʲo][14] 'you are tall'

(43) a. /sióx/ [sʲóx] 'fish'
b. /síox/ [síox] 'deceive'

The stress in (42b-43b) blocks Pal from applying.

Similar examples can be found in Mengen showing stress blocking LHS as well as Lab and Pal. Rath (1981:24) indicates stress is nonphonemic, although the following minimal pair differentiated only by stress is noted.

(44) a. /ílo/ 'fish species'
b. /ilό/ 'type of dance'

More recently, Rath (personal communication) has found more extensive stress differences, and so now analyses stress as contrastive. This contrastive stress differentiates syllabic from nonsyllabic vocoids in (45-46).

(45) a. /mué/ [mʷé] 'snake'
b. /soáli/ [sʷáli] 'bad'
c. /gúa/ [gú$a] 'ginger'
d. /bóalala/ [bó$alala] 'taro species'

(46) a. /táina/ [táʲ$na] 'banana species'
b. /taína/ [ta$í$nʌ] 'taro species'

Lab applies in (45a-b), resulting in the initial consonant coalescing with the following back vowel. Stress blocks Lab in (45c-d), however, so the back vowels remain syllabic and form a separate syllable. Similarly, LHS applies in (46a), resulting in the diphthong [aʲ]. Once again,

stress on the high vowel in (46b) blocks LHS so the /i/ forms a separate syllable.

The forms in (46) are relevant to the arguments presented in section 2.2 regarding the status of diphthongs such as [a^{i}]. It was argued there that the two parts of the diphthong were in some sense equal. They are not, however, equal in regard to stress. While stress on the high vocoid will block LHS, stress on the low vocoid will not. This suggests that in some way it may be the high vocoid that is made nonsyllabic by LHS.

In summary, a stressed vocoid will be syllabic even if it occurs in an environment in which the unmarked value for the vocoid is [-syll]. It should be noted that no instances have been found of stress blocking WIN or PVN. Whether this is significant or accidental is not clear.

3.2 Morpheme breaks

In many instances, sequences across morpheme boundaries syllabify in the same way as do sequences within a morpheme. Seiler (1985:18) uses this fact in Imonda to justify his analysis of diphthongs as phonemic vowel sequences. Notice the following examples.

(47) /kau/ 'chin'

(48) /sue la-u/ 'light a fire!'
fire light-IMP

(49) /ka-m fa-ai-h-u/ 'give me!'
1-GL CL-give-REC-IMP

Seiler notes that while the /u/ in (48-49) represents the same morpheme, it is syllabic in (49), but nonsyllabic in (48) just as it is nonsyllabic in (47). It is clear that in Imonda LHS applies both within morphemes and across morpheme boundaries.

In other instances, however, a morpheme boundary does seem to block nonsyllabification. Lichtenberk (1980:14-15) formalizes three rules in which vowels are desyllabified word initially or postvocalically when before another vowel. All three, then, are language specific realizations of the processes of WIN and PVN. In each case there is a condition that "no morpheme boundary intervenes between the focus [the vowel to be nonsyllabified - JMC] and the following (vowel)." Clifton (1987:8) notes a similar constraint on WIN in Kope.

The interplay between morpheme boundaries and the nonsyllabification processes raises interesting questions in the analysis of Selepet vocoid sequences. As noted in section 2.2, the syllabification of most vocoid sequences is predictable in Selepet in terms of LHS and BFS. McElhanon (1967) does, however, find contrasts in syllabification for certain sequences as shown in (50-51).

(50) a. [ⁱek] 'look at them!'
b. [si$ep si$ep] 'kind of bird'

(51) a. [ⁱɔk] 'he'
b. [i$ɔk] 'let him sleep!'

In McElhanon's analysis, forms (50b-51b) are the expected forms since he assumes all sequences of V_1V_2 should syllabify as two syllables when V_1 is higher than V_2. Therefore, McElhanon claims the nonsyllabic vocoids in (50a-51a) are entered into the lexicon as semivowels.

If, however, WIN and PVN also apply in Selepet, (50a-51a) are natural forms since the high vowel is in word initial position. Form (50b) is also natural, since the high vowel is not in word initial position and, as noted above, V_1 is higher than V_2. It is form (51b) that is apparently problematic since the word initial high vocoid is syllabic. It may be possible, however, to account for this form in terms of morpheme boundaries. According to McElhanon (1970:21), /+ɔk/ is the morpheme for the 'inchoative future' ('let him ...!'). If we assume WIN is blocked in Selepet by a morpheme boundary, an underlying morphemic structure of /i+ɔk/ would result in (51b).

The claim that a morpheme boundary blocks WIN does not mean it blocks all the nonsyllabification processes outlined above. In particular, forms like (52) indicate that boundaries do not block LHS.

(52) [gaᵉk] 'let him come!'

McElhanon (1970:21) indicates that the 'inchoative future' is /+ɔk/ following high and mid vowels, but /+ek/ following low vowels. In (52) the sequence /a+e/ is realized as a single syllable even though it occurs across a morpheme boundary.

The discussion of Selepet points out a difference between stress and morpheme boundaries. As noted in section 3.1, stress seems to block LHS, Lab, and Pal, but not WIN or PVN. Morpheme boundaries, on the other hand, seem to block WIN and PVN but not LHS, Lab, or Pal. Once again, it is unclear whether this difference is accidental or not.

3.3 Long vowels

Long vowels do not actually block nonsyllabification. They may, however, set up surface contrasts in syllabification. Clifton (1987a:8) gives the following forms from Kope which seem to show contrastive syllabification.

(53) a. [ᵘe] 'arrow type'
b. [u.e] 'sugar cane'

As the vast majority of syllabification is predictable, however, these contrasts are suspect. In accounting for the contrasts Clifton

(1987a:10) gives the following examples showing vowel length is contrastive in Kope.

(54) a. [abea] 'father'
b. [abeea] 'sago squeezing bag'

(55) a. [obo] 'water'
b. [oobo] 'woman'

(56) a. [tutu] 'handle'
b. [tutuu] 'long'

Given this contrast between long and short vowels, Clifton (1987a:13) analyses the contrasts in (53) in terms of vowel length as shown in (57).

(57) a. /ue/ [ᵘe] 'arrow type'
b. /uue/ [u$ᵘe] 'sugar cane'

In (57a) the initial /u/ is nonsyllabic due to WIN, resulting in a monosyllabic form. In (57b), on the other hand, the second /u/ is nonsyllabic due to PVN. This, in turn, blocks WIN from applying to the initial /u/. Thus, a surface contrast in syllabification can be attributed to an underlying contrast in vowel length.

4 Conclusion

Mohanan's claim that the unmarked state for all vocoids is [+syll], that is, syllable peaks can best be viewed as a paradigmatic condition. That is, in the absence of environmental conditioning, a vocoid will be syllabic. In the case of sequences of vocoids, however, it seems syntagmatic factors frequently dictate that they are syllabified as less than the maximum number of syllable peaks. In these environments the paradigmatically preferred state is seen only when the 'nonsyllabification' of vocoids is blocked by various conditions such as stress or morpheme breaks.

NOTES

* Many of the ideas for this paper were first presented in Clifton (1987b), and were further developed in workshops in the Papua New Guinea branch and the Irian Jaya program of the Summer Institute of Linguistics. Special thanks go to Dottie James, Eileen Gasaway, and Duane Clouse for input.

1. I will use the feature [-syll] to refer to segments which Clements and Keyser (1983) would have attached to a C node, and the feature [+syll] to refer to segments which they would have attached to a V node.

2. In phonetic transcription, raised vowel symbols will generally be used to indicate nonsyllabic vocoids. In phonemic transcription, nonsyllabic vocoids will generally be represented by the semivowels y and w before another vocoid, and by normal vowel symbols after another vocoid. Syllable boundaries will only be shown when relevant to the discussion.

3. The term 'process' is not used here as a theoretical concept. What is referred to as a process may be a morpheme structure constraint, a lexical rule that fills in unspecified features (Kiparsky 1982:59-65), or part of some 'principles of headship' (Mohanan 1986:30).

4. WIN and PVN are combined as a single process in Clifton (1987a:8).

5. Tone has been omitted from the examples as it is not relevant to the discussion.

6. In a language with closed syllables, we would also expect high vocoids to be nonsyllabic before a consonant following a vocoid. As will be shown in section 2.2, however, high vocoids following another vocoid are generally nonsyllabic even in languages which do not permit closed syllables. Thus, there is no need for a separate process to account for the nonsyllabicity in this environment.

7. McElhanon (1967:9) makes it clear that the claim that the sequence [a^{u}] in [ba^{u}] 'pig' is a complex (i.e. close-knit) nucleus does not mean the syllable structure is simply CV. Instead, McElhanon's inventory of syllables includes VV, VVC, CVV, and CVVC. If LHS is formalized within CV phonology (Clements and Keyser 1983), the new structure would be a branching V node.

8. The situation regarding front vowels is unclear. Examples like /eala/ 'that one' and /lea/ 'fish' are given in which it is unclear whether or not the /e/ is syllabic, and whether or not it merges with /i/.

9. Strictly speaking, these claims imply that the phonetic transcriptions above are incorrect since they show the second vowel as nonsyllabic. In section 3.1, however, evidence is presented that the high vowel is in some way subordinate.

10. There is some question as to whether each vowel is a separate mora phonetically or whether each vocoid must be counted for processes like stress assignment whether or not they are separate mora. If it is the second, it simply means LHS may in fact apply in Buin.

11. This implies that LHS must be sensitive to syllable structure created by PVN.

12. It will be shown in section 2.3 that the /i/ in [k^{i}u] is realized as palatalization on the /k/. In a reduplicated prefix, however, it is syllabic.

13. Seiler gives the phonemic form of (30) as /huef/, but indicates that /h/ generally represents a velar fricative.

14. The [^{i}] in (42b-43b) is transitional according to Brown.

REFERENCES

Anderson, S.R. 1985. Phonology in the twentieth century. Chicago: University of Chicago Press.

Brown, B. 1987. A quantitative phonology of Mai Brat. ms, SIL, Jayapura, Irian Jaya, Indonesia.

Chomsky, N. and M. Halle. 1968. The sound pattern of English. New York: Harper & Row.

Clements, G.N. and S.J. Keyser. 1983. CV phonology: A generative theory of the syllable. Cambridge, MA:MIT Press.

Clifton, J.M. 1987a. Orthographic reform in Kope, Studies in Melanesian orthographies, ed. by J.M. Clifton, Datapapers on Papua New Guinea languages, vol. 33:1-15. Ukarumpa via Lae, Papua New Guinea: SIL.

_____. 1987b. Vowel sequences in CV phonology, paper presented to Linguistic Society of Papua New Guinea.

_____. in progress. A phonological sketch of Kope.

Gasaway, E. and H. Sims. 1977. A preliminary phonology of Girawa. ms, SIL, Ukarumpa via Lae, Papua New Guinea.

Griffin, M. 1986. Writing tok ples Terei. ms, SIL, Ukarumpa via Lae, Papua New Guinea.

Hynum, D. 1987. Numanggang phonology. ms, SIL, Ukarumpa via Lae, Papua New Guinea.

Kiparsky, P. 1982. Lexical morphology and phonology, Linguistics in the morning calm, ed. by Linguistic Society of Korea. Seoul: Hanshin.

Lichtenberk, F. 1980. A grammar of Manam. Ph.D. dissertation, University of Hawaii.

McElhanon, K.A. 1967. Selepet vocoid clusters. Papers in New Guinea linguistics No. 6, Pacific Linguistics A-12:1-18.

_____. 1970. Selepet verb morphology. Papers in New Guinea linguistics No. 10, Pacific Linguistics A-25:19-35.

Mohanan, K.P. 1986. The theory of lexical phonology. Dordrecht: D. Reidel Publishing Co.

Pike, K.L. 1947. Phonemics: A technique for reducing languages to writing. Ann Arbor, MI: The University of Michigan Press.

Rath, D. 1981. Mengen phonology: a generative view. ms, SIL, Ukarumpa via Lae, Papua New Guinea.

Seiler, W. 1985. Imonda, A Papuan language. Pacific Linguistics B-93.

Smith, J. and P. Smith. 1974. Mianmin phonemes and tonemes. Studies in languages of the Ok family, Workpapers in Papua New Guinea languages, vol. 7:5-33.

Turner, B. 1986. Manam teaching grammar. Datapapers on Papua New Guinea languages, vol. 34. SIL: Ukarumpa via Lae, Papua New Guinea.

PRAGMATICS AND GRAMMAR: MOTIVATION AND CONTROL

Robert A. Dooley

1 Introduction
2 Pragmatics in relation to semantics and syntax
  2.1 Early definitions of pragmatics
  2.2 The interface between pragmatics and semantics
  2.3 The interface between pragmatics and syntax
  2.4 Externally motivated phenomena encoded by grammar
  2.5 Types of interaction between grammar and pragmatics
3 Pragmatic phenomena with little or no grammaticalization
  3.1 Mbya hearsay particle
  3.2 Wayampi main clause word order
4 Pragmatic phenomena with partial grammaticalization
  4.1 Mbya future marker
  4.2 Hixkaryana hearsay particle
  4.3 Hixkaryana afterthought elements
5 Pragmatic phenomena with more or less complete grammaticalization
  5.1 English WH-clefts
  5.2 Dutch auxiliary placement
6 Concluding remarks

1 Introduction

Much of what has been done as "pragmatics" has dealt with utterance interpretation, that is, with the influence of context on meaning. Relatively little has been said about the interaction between context and form. As a result, there is much we do not know about how pragmatic phenomena relate to the grammar of a language. The present paper makes three points in this regard: (a) it is not anomalous for there to be pragmatic phenomena encoded in the syntax and controlled by the grammar; (b) it is not anomalous for there to be pragmatic phenomena encoded in the syntax which show little or no grammatical control; and in general, (c) for pragmatic phenomena encoded in the syntax, grammatical control can be present in different degrees and in different ways.

Point (a) is not particularly controversial in principle, as the brief survey of the status of pragmatics and pragmatic phenomena in current linguistic theory (Sect. 2) should make clear. Point (b) is illustrated in Sect. 3, and (a) and (c) in Sects. 4 and 5. Examples are drawn primarily, but not exclusively, from languages of lowland South America. The phenomena under discussion are ordered according to roughly increasing degrees of grammaticalization. Sect. 6 provides a concluding discussion.

This paper, then, is of the genre that deals with external explanations for certain syntactic facts. Hyman (1984:67) gives the distinction between internal and external explanations as follows: "an internal explanation will propose an account in terms of the nature of syntax itself, while an external explanation will attempt to relate the syntactic problem to phenomena outside the realm of syntax (e.g. semantics or pragmatics)". But even though the focus of this study is on external explanations, it assumes that there are syntactic phenomena which require internal explanations, and that such explanations constitute the realm of grammar. The point being made is that along with internal explanations, or in some cases in lieu of them, there is a need for external explanations. Whether or not one deals with external explanations depends in part on the kind of linguistic description he is interested in, and in part on the kind of phenomena he chooses to examine.

2 Pragmatics in relation to semantics and syntax

Pragmatics is notoriously hard to define. There are at least four main problems. One is that there are many different ways to slice the linguistic pie into such things as pragmatics, semantics, and syntax, and each person tends to have his own way of doing it. A second difficulty is that pragmatics is often defined in terms of adjacent "pieces" of the pie (semantics and morphosyntax), and there is no general agreement on the domain of these adjacent pieces. The third problem in coming up with a neat linguistic distinction is that there may be no corresponding neat distinction in the reality of language; the interface between pragmatics and other subareas of language may be intractably fuzzy. And the fourth difficulty is the fact that, since so many different things have been discussed under the rubric of pragmatics, it may not constitute a unified or natural field of study at all; see, for example, discussion in Verschueren 1987. Nevertheless, there is general consensus that pragmatics deals with what John Gumperz calls "contextualization phenomena" (Verschueren 1987:26), different ways in which the context of an expression interacts with its form or interpretation.

2.1 Early definitions of pragmatics

Early pioneers in defining pragmatics were Charles Morris and Rudolf Carnap. Their strategy, which in broad outline has been followed up to the present, was to distinguish it from syntax on one hand and semantics on the other. The following is taken from the introduction of Searle, Kiefer, & Bierwisch 1980:

> According to Morris's earliest formulation of this distinction (1938), syntactics studies "the formal relations of signs to one another". Semantics studies "the relations of signs to the objects to which the signs are applicable". And pragmatics studies "the relations of signs to interpreters". But this

> distinction between pragmatics and semantics is very unsatisfactory. For example, taken strictly, the above definitions would have the consequence that pragmatics is a branch of semantics, since signs are clearly "applicable" to interpreters. Morris later modified this definition, and redefined pragmatics as "that branch of semiotics which studies the origins, the uses, and the effects of signs" (1946). Carnap (1942), following Morris's earlier position, gave the following definition, which has proved influential to subsequent authors: "If, in an investigation, explicit reference is made to the speaker, or to put it in more general terms, to the user of the language, then we assign it to the field of pragmatics... If we abstract from the user of the language and analyze only the expressions and their designata, we are in the field of semantics. And if, finally, we abstract from the designata also and analyze only the relations between the expressions, we are in [logical] syntax. The whole science of language, consisting of the three parts mentioned, is called 'semiotics'" (p. viii).

2.2 The interface between pragmatics and semantics

Many subsequent treatments of pragmatics have been largely concerned with the interface between pragmatics and semantics. Searle, Kiefer, & Bierwisch state three conceptualizations of pragmatics which differ according to which view of semantics is being assumed (pp. ix-xi). First, against a background of formal philosophy and logic in which an expression's interpretation is given in terms of entities it denotes in some world, pragmatics is concerned with how an expression's interpretation can depend on the particular context of its use. This point of view is represented "classically" by Carnap and more recently by Stalnaker (1972). Second, against a semantic theory in which an expression can have both a context-free, literal meaning (its sense) and a context-dependent meaning, pragmatics is concerned with how speakers and hearers arrive at context-dependent meanings. Katz (1977) and Sperber & Wilson (1986) represent this viewpoint. Third, in the tradition of Austin (1962), Grice (1968, 1975, 1978), and Searle (1969, 1975, 1976), where the focus is upon speech acts and illocutionary force, the idea of context-free meaning is fictitious. On this view there is no semantics apart from pragmatics; all meaning involves the use of contextual conditions. This is "radical pragmatics" of a rather absolute type (Cole 1981).

The above sketch illustrates that, in many treatments of pragmatics, the focus has been very much on meaning and interpretation -- the interface of pragmatics with semantics. Thus, pragmatics has been defined as "meaning minus truth conditions" (Gazdar 1979:2), "the study of meaning in relation to speech situations" (Leech 1983:6), "the study of the interpretation of utterances" (Sperber & Wilson 1986:10), and "meaning minus semantics" (Roland Posner, cited in Verschueren

1987:14). This focus is not surprising, given the many interesting and important questions in this area, and given the interest in illocutionary force and utterance interpretation that has given impetus to the study of pragmatics from Austin 1962 until the present. Nevertheless, during this period the interface between pragmatics and syntax has generally been overlooked.

2.3 The interface between pragmatics and syntax

I follow Fillmore in a view of pragmatics which allows for an active interface with syntax as well as with semantics: "Syntax... characterizes the grammatical forms that occur in a language, whereas semantics pairs these forms with their potential communicative functions. Pragmatics is concerned with the three-termed relation that unites (a) linguistic form and (b) the communicative functions that these forms are capable of serving, with (c) the contexts or settings in which those linguistic forms can have those communicative functions. Diagrammatically,

Syntax [form]
Semantics [form, function]
Pragmatics [form, function, setting]"
(Fillmore 1981:144).

While this formulation requires some amplification (the phrase "communicative function", for example), it clearly provides for an interface between pragmatics and syntax, since it includes linguistic form as one of the factors with which pragmatics deals. Fillmore has more recently made reference to this kind of interface: "The proposal I want to support is that there are many linguistic structures that appear to be dedicated to specific pragmatic purposes... there are a lot of grammatical constructions, there are lots of lexical items in particular grammatical contexts, that require for their interpretation an anchoring in some kind of a real situation; and beyond that there is a whole lot of pragmatics that is not connected with grammar" (Verschueren 1987:18, 43). That is, "some syntactic facts require semantic and pragmatic explanations and... some semantic facts require pragmatic explanations" (Fillmore, 144).

The notion of context, or setting as Fillmore refers to it, is generally recognized as being central to pragmatics. Gumperz (Verschueren 1987:26), in fact, prefers the term "contextualization phenomena" to "pragmatic phenomena". As Parrett (op. cit., 12) points out, "context is a very vague concept... there are existential contexts, psychological contexts, there are situational contexts", as well as linguistic contexts. Furthermore, as Gumperz notes, "there are some people who look at context as pre-existing the interaction. There are some people who look at context as emerging in the interaction. Some people think both views are valid" (op. cit., 47). For the purposes of this paper, the relevant context of a given expression in a given

instance is understood to comprise those factors which are external to it and interact with its form or interpretation (cf. Leech 1983:13).

2.4 Externally motivated phenomena encoded by grammar

The interface between pragmatics and syntax raises the question as to whether or not formal grammatical structures can be motivated by language-external, real-time phenomena. By "real-time" I am referring to discourse and other phenomena which motivate linguistic usage in particular instances, not merely those which have operated diachronically. That is, here my focus is not on "the functional explanation [which] applies on the evolutionary level -- either the evolution of the organism or the language" (Chomsky 1980:23). In regard to external, real-time phenomena, two extreme positions may be stated:

(a) formal grammatical structures are NEVER a result of real-time externally motivated phenomena;

(b) ALL structural phenomena in language are the result of real-time externally motivated phenomena.

Both positions (a) and (b), although they had in the past, and may still have adherents, are largely straw men in current linguistic debate. Against position (a), Chomsky (1975:56ff) states: "Surely there are significant connections between structure and function; this is not and never has been in doubt... Where it can be shown that structures serve a particular function, that is a valuable discovery." One example of Chomsky's own work in this regard is his study of focus-presupposition phenomena (Chomsky 1971:200ff). Newmeyer (1983:11f) makes a similar statement: "No generativist denies the interest of determining the discourse function (if any) of a particular syntactic construction or constraint... No generativist, to my knowledge, has ever disparaged the study of the interaction of form and function." Against (b), Givon (1979:82) states: "Rather than wind up with a formal and autonomous level of structural organization in language, we do indeed find syntax to be a DEPENDENT, functionally motivated entity whose formal properties reflect -- perhaps not completely, but nearly so -- the properties of the explanatory parameters that motivate its rise." Of interest here are the words "perhaps not completely" and Givon's recourse to non-real-time explanatory parameters.

The difference, then, between formalists and functionalists is not that one is at the opposite end of the spectrum from the other; nor does the difference turn on whether externally motivated phenomena exert historical influence on linguistic structure. The difference, rather, is that generativists see the influence of externally motivated, real-time phenomena as slight, while functionalists see it as more or less pervasive. Speaking of the system of grammatical rules for a language, Chomsky (1982:115) says, "The thesis of autonomy of syntax says that this system is pretty well self-contained. Undoubtedly, the system

interacts with other systems, but the interaction is more or less at the periphery." Similarly, Newmeyer (1983:111): "though we might find correlations between grammatical structures and their discourse functions, we only rarely find direct reflections of the latter by the former." The position of Givon, towards the other end of the scale, has already been cited. The question of whether the influence of such factors is at the periphery or the core of language is an empirical one which at the present has no generally accepted answer.

Thus, while not minimizing the important differences between generative and functionalist paradigms, I simply note here that even in the formal framework, it is in principle not anomalous for certain pragmatic phenomena to be grammaticalized.

2.5 Types of interaction between grammar and pragmatics

A useful distinction can be drawn between grammar and (morpho)syntax. I take the term "grammar" to refer to the body of system-internal rules which controls large portions of morphosyntax, phonology, and perhaps semantics. Thus grammar is, by definition, taken to deal with internal explanations in Hyman's (1984) sense. With this understanding, I will illustrate here that not all of morphosyntax need be controlled by grammar; a certain amount may be controlled directly from syntax-external considerations, such as from the area of discourse-pragmatics. Another way to state this would be to say that not all syntactic regularity or all rules of syntax, need be grammatical. To the extent that this is so, we use "autonomous" as a restrictive modifier when we speak of autonomous (morpho)syntax. Extralinguistic factors can influence morphosyntax both through the mediation of grammar, as noted by Chomsky and Newmeyer, or more directly, as will be illustrated, in lieu of grammar. In regard to a given syntactic phenomenon, it is always appropriate to ask whether it is best described grammatically or by an external explanation; in many cases, it is profitable to identify aspects of a phenomenon which lend themselves to one kind of description, and others to a different sort.

In the following sections I examine several different pragmatic phenomena and their interactions with grammar. By "pragmatic phenomena" I have in mind morphosyntactic phenomena which reflect real-time, pragmatic considerations; specifically, I will be dealing with motivation generally referred to as "discourse-pragmatic". Certain of these phenomena (Sect. 3) involve no apparent grammaticalization; they affect morphosyntax directly, without appreciable recourse to the mediation of grammar. Other pragmatic phenomena (Sect. 4) are externally (pragmatically) controlled part of the time, and internally (grammatically) controlled part of the time. Still others (Sect. 5) are fully grammaticalized, even to the point (Sect. 5.2) where real-time original pragmatic motivation cannot be consistently maintained.

3 Pragmatic phenomena with little or no grammaticalization

Here I illustrate the claim that not all of syntax is under grammatical control, and specifically that there exist pragmatic phenomena with little or no apparent grammatical control. Instances of this are predictably hard to come by, since "the driving force of grammar is to get control of whatever it can" (Hyman 1984:80); grammar tends to act as if all of syntax lay in its rightful sphere of influence. This can perhaps be understood in light of the natural human tendency to conventionalize, and even prescribe, things which are done frequently.

The illustrations in this section are from two Brazilian languages of the Tupi-Guarani family: Mbya Guarani and Wayampi.

3.1 Mbya hearsay particle

In Mbya, the basic word order is SVO and the basic or neutral pragmatic configuration is Topic-Comment. There exist other, special-purpose configurations which are used only under certain specifiable discourse conditions; these include Focus-Presupposition and a type of Topic-Comment configuration in which the Topic is given phonological salience. For more on pragmatic configurations in Mbya, the reader is referred to Dooley 1982, 1987a.

The division of an utterance into components such as Topic, Comment, Focus, and Presupposition is signalled by a variety of means. Among these are boundary phenomena such as an intonation break, pause, and the occurrence of certain particles which function as "spacers". Spacers are typically unstressed clitics of one or two syllables; semantically, they are sentential-scope operators signalling such things as mood (including evidentiality), tense, or aspect. Their positioning within the sentence is a pragmatic phenomenon: they are placed so as to indicate a constituent boundary in the pragmatic configuration of an utterance (Dooley 1982; 1987a; 1987b, Sect. 4.2). Their effect as spacers is due to their positioning plus their phonological and semantic properties, all of which combine to suggest a break in processing between message units (Osgood & Sebeok 1954:52, Haiman 1985:105).

One such spacer particle is **je** 'hearsay'. As with other Mbya spacers, it cliticizes to some grammatical phrase which is a clause constituent. Granting this grammatical control on the phrase level, the evidence I will present shows that the sentence positioning of **je** -- the fact that it attaches to one expression rather than to another -- is due to discourse-pragmatic factors rather than any apparent rule of grammar. For an overall view of pragmatic configurations in Mbya, the reader is referred to Dooley 1982.

In 1, **je** separates Topic from Comment:

(1) Context: A hawk (unrecognized as such) was buying chickens from local farmers. He told the farmers that he would come to get them one by one, and that he didn't want them enclosed. He left, and afterwards when other farmers asked what had transpired, the farmers who had talked with the hawk said:

```
-Topic-     -------Comment-------
uru     je  nha-mboty   eme
chicken HSY 1+2-enclose NEG.IMPER
'In regard to the chickens, as he said, let's not
enclose them.'
```

Ex. 1 shows a fronted direct object **uru** 'chicken' as Topic, here a resumptive topic of conversation. The particle **je** occurs "in the cracks" between Topic and Comment.

In 2 **je**, in combination with the aspect enclitic **tema** 'continuously', occurs between Focus and Presupposition:

(2) Context: A group of Mbya were on a long journey.

```
---Focus----          --Presupposition-
ka'aguy anho tema je  o-axa  o-je'oi-vy
woods   only CONT HSY 3-pass 3-go.PL-SER
'They were only going through woods' [in contrast
to open roads and fields].
```

In 3, **je** follows a sentence-initial connective, which is another type of pragmatic component in Mbya:

(3) Context: The local "chief of police" had sent his men out to lie in wait at every crossroads for a certain wanted man.

```
Connective-     Topic ------Comment-----
ha'e rire  je   Ø     o-aró  o-kua-py
3.ANA after HSY Ø     3-wait 3-be.PL-SER
'After that, they all waited for him.'
```

In 3, the subject 'they', represented by zero anaphora, is treated as an unmarked (not especially informative) Topic.

For the examples presented thus far, one could posit not one but two possible grammatical rules for the sentence positioning of **je**: either in second position in the sentence, or preceding the main verb. It turns out, however, that Mbya has no items whose positioning can be

described in either of those ways, and the following examples show that something further is happening with **je**. In 4, it occurs after a Setting expression which is itself in second position:

(4) Context: A woman and her brother, in danger, were having to flee to another location.

```
Connective- --Setting--      Topic -Comment-
ha'e  rire  o-vaé     rá je  Ø      o-porandu
3.ANA after 3-arrive DS HSY Ø      3-ask
'Afterwards when they had arrived, they [the people
of that place] asked: ["Why have you come?"]' (7:29)
```

In 5, **je** occurs both after an initial connective 'after that' and between Focus and Presupposition:

(5) Context: The preceding paragraph spoke of a group of Mbya travelling for one day. The following sentence begins a new paragraph:

```
Connective-      ----------Focus---------
ha'e  rire  je  peteĩ jaxy ha'e  javi re  je
3.ANA after HSY one   moon 3.ANA all  for HSY

---Presupposition--
o-guata  o-je'oi-vy
3-travel 3-go.PL-SER

'After that, it was for a whole month that they
travelled.'
```

6 has four occurrences of **je**:

(6) Context: Text-initial sentence: 'Once a certain man went to the woods and saw a lot of wild game animals.' Second sentence:

```
--Conn--      ----------Setting----------
ha'e  vy je  o-juka ta       o-iko-vy jave je
3.ANA SS HSY 3-kill about.to 3-be-SER when HSY

----------Setting-----------      --Setting--
peteĩ ava o-vaé    ha'e  py vy je  ij-ayvu  vy je
one   man 3-arrive 3.ANA in SS HSY 3-speech SS HSY

Topic ------Comment------
Ø     aipo-e'i   ix-upe
ANA   thus-3.say 3-DAT
```

'And so, just as he was about to kill them, a certain man arrived there, and he [the second man] spoke up and said to him: ["Don't kill them."]'

Note that the second occurrence of **je** in 6 is not juxtaposed to any verb. There does not appear to be any theoretical limit to the number of occurrences of **je** in a sentence, but all occurrences take place at boundaries in a plausible pragmatic configuration, plausible with respect to the flow of information in that context. Of the grammatical rules which have been found for particle positioning in lowland South American languages (Dooley 1987b), none account for the the placement of **je** in Mbya. It can be concluded that its placement in the sentence is a pragmatic phenomenon which shows no apparent grammaticalization: it is due to the direct influence of pragmatic motivation on the syntax.

3.2 Wayampi main clause word order

Thompson (1978) makes the point that "some languages utilize predicate-argument order primarily for pragmatic purposes, and some primarily for grammatical purposes. There are also languages which use predicate-argument order for both purpose[s] without giving priority to either" (20). She cites Russian and Mandarin Chinese as languages with basically pragmatic word order; in those languages, "the known information tends to be placed sentence initially and the new information sentence finally" (21). In such languages, the overall ordering of sentence elements is largely a pragmatic phenomenon, not controlled by the grammar.

A language whose word order has come under study for similar reasons is Wayampi of northern Brazil (Jensen 1980, Payne 1987). In Wayampi, however, pragmatic categories are manifested in word order in a different way: in main clauses, items which the speaker is presenting as particularly new or informative precede the verb, while less informative items follow it. More specifically, "items precede the verb when:

(a) the speaker instructs the hearer to establish a new cognitive file, or entry, for the item, relative to the universe of discourse;
(b) the speaker instructs the hearer to establish the item as a major topic for the ensuing discourse, in contrast to whatever has been the major topic up to that point;
(c) an already-established major topic will precede the verb if it occurs in a restatement clause which closes a thematic discourse unit...;
(d) the information is highly pragmatically marked.

Items otherwise follow the verb when:

(e) the hearer is not instructed to establish a new cognitive file, either because such a file has already been established, or because the speaker anticipates it will not be needed. Specifically included are items which the speaker assumes are already

identifiable to the hearer, and some non-referential entities" (Payne 1987).

The influence of pragmatics on main clause word order in Wayampi is illustrated in 7 by a portion of a text describing various kinds of birds (Jensen 1980):

(7) a. mytõ ywy-rupi-te oata S V
mutum ground-along-only 3.walk
'The mutum (currasow) walks only along the ground.'

b. oke-ta reme-te wate ojupi V
3.sleep-FUT when-only above 3.climb
'Only when it is going to sleep, it climbs up above.'

c. owewe wate V
3.fly above
'It flies above.'

d. wate-te mytõ oke S V
above-only mutum 3.sleep
'The mutum sleeps only up above.'

e. ywyra 'a o'u mytõ O V S
tree fruit 3.eat mutum
'The mutum eats fruit from trees.'

f. ka'a ro o'u mytõ O V S
jungle leaf 3.eat mutum
The mutum eats jungle leaves.'

g. pira o'u O V
fish 3.eat
'It eats fish.'

h. mytõ o'u S V
mutum 3.eat
'The mutum eats them.'

i. jakami iwowe S V
jacamim likewise
'The jacamim is the same way.'

Sentences 7a-h constitute a paragraph about the mutum, and 7i begins a new one about the jacamim. (In the indication of the order of the major clause constituents that is given to the right of each example, the one referring to the mutum is in boldface.) The preverbal position of **mytõ** 'mutum' in 7a indicates the speaker's instructing the hearer to establish a new cognitive file for it (Payne's point (a) above), and

also to establish it as a major topic for the ensuing part of the discourse (point (b)). Once so established, it reverts to postverbal position (point (e)) or nonphrasal representation in 7b-g. 7d appears to be an exception to this, but it could also be an instance of closing a small thematic unit (point (c)), one consisting of the three sentences 7b-d and dealing with where the mutum sleeps. In 7h we find another preverbal occurrence at the end of perhaps two coterminous thematic units: 7a-h about the mutum in its different habits, and 7e-h specifically about what it eats. In 7i, where begins a thematic unit about the jacamim, **jakami** is found in preverbal position. In 7e-g, preverbal position is accorded to the three items which the mutum eats: **ywyra 'a** 'fruit from trees', **ka'a ro** 'jungle leaves', and **pira** 'fish'. These three items can be considered "highly pragmatically marked" (point (d)) in the sense that they constitute the most informative elements of their respective utterances.

Other examples of highly pragmatically marked information in preverbal position include contrast (ex. 8a-b) and the answers to information questions (9b) (each of these two examples is composed of contiguous text material; Jensen 1980):

(8) a. jimi'apuku jaopy
long.flute we.play
'We play a long flute.'

b. takwari papa opy
small.flute father plays
'Father plays (by contrast) a small flute.'

(9) a. mope-kyty wyi poko owae uu?
where-place from Q 3.arrive come
'Where did he come from?'

b. kyty wyi te uu
place from only come
'He came from right there.'

Although a full analysis of word order in Wayampi is not currently available, and there is some evidence that SVO or (particularly) SOV, or both, may have a more basic status than other orders (Jensen 1980), facts such as those presented above indicate that discourse-pragmatics is the major motivating factor for word order in main clauses. Subordinate clauses are consistently (S)OV (ibid.).

4 Pragmatic phenomena with partial grammaticalization

The linguistic picture would be simpler if pragmatics and grammar acted separately and independently upon morphosyntax. However, such is not always the case. Given a view of grammar as a type of conventionalization (Givon 1979, Dooley 1983), it is not surprising that

we find pragmatic phenomena which are grammaticalized in different ways and to differing degrees. This section provides three illustrations of partial grammaticalization of pragmatic phenomena. Sect. 5 illustrates grammaticalization of a more complete kind.

4.1 Mbya future marker

One way that grammaticalization can be seen as partial is by operating on a part-time basis. That is, something may be a pragmatic phenomenon and under more or less complete pragmatic control part of the time, and a nonpragmatic phenomenon under grammatical control the rest of the time. This is the case with the positioning of the Mbya Guarani future tense marker **va'erá** and its contracted form **'rá**.

In Mbya, **va'erá** occurs in the verb phrase following the main verb:

(10) ko'ẽ rá ja-vy'a va'e-rá ja-kua-py
dawn DS 1+2-be.happy REL-FUT 1+2-be.PL-SER
'Tomorrow we will celebrate together.'

When the speaker is expressing strong interpersonal feeling -- lack of patience, for example -- the contracted monosyllabic form **'rá** can be used instead:

(11) Ne-kane'õ vy nd-ere-o-ve-i 'rá.
2SG-tired SS NEG-2SG-go-more-NEG FUT
'When you get tired, you won't go on any farther'
(said in a disparaging way).

(12) Ha'e rami teĩ xee a-a tema 'rá.
3.ANA like but 1SG 1SG-go CONT FUT
'Even so, I'll keep right on going'
(said in reply to an utterance like 11).

Both **va'erá** and **'rá** are unstressed enclitics.

When **va'erá** and **'rá** occur in the verb phrase after the main verb, they are occurring in a position which is grammatically determined and can be described in terms of grammatical categories. Consider, however, the sentences of 13:

(13) Context: A young lady was to throw a flower to indicate her choice of a husband. Her sisters were to do the same. At different points in the text, she told both her mother and her sisters which young man was her choice:

a. Topic (----------Focus--------- --Presup-)
xee ko pe-va'e ae re rive 'rá a-mombo
1SG OP nearby-REL EXA at merely FUT 1SG-throw
'As for me, I'll throw [mine] merely at that one.'

b. -------Focus-------- (Top (Topic -Comment-))
pe-va'e re rive 'rá ko xee yvoty a-mombo
nearby-REL at merely FUT OP 1SG flower 1SG-throw
'It's merely at that one that I will throw my flower.'

In neither 13a nor 13b does the future marker **'rá** occur in its grammatically determined postverbal position. Rather, it is positioned as a spacer to help indicate constituent boundaries in the pragmatic configurations that are labelled. The enclitic **ko**, which indicates that the speaker recognizes that what is being said reflects a personal viewpoint, is also used as a spacer. The difference between **'rá** on one hand and **ko** and **je** (Sect. 3.1) on the other, is that whereas the latter two are always spacers, the task of positioning **'rá** is split between grammar and pragmatics. There is a grammatically determined position in which it occurs when it is not particularly needed as a spacer; but when pressed into service by pragmatics, it functions as a spacer. When the Mbya future marker is used in this way, it is always the contracted form **'rá** which occurs. As a result, when **'rá** is used as a spacer, it is not capable of encoding the nuance of strong feeling that it does when in paradigmatic opposition with the longer form **va'erá** in postverbal position. In 14, another sentence from the same text as 13 but spoken by one of the sisters, the future marker **'rá** is not employed for pragmatic ends, hence occurs in its grammatically determined position:

(14) ----Focus---- (Top ---Comment---)
pe-va'e re ko xee a-mombo 'rá
nearby-REL at OP 1SG 1SG-throw FUT
'It's at that one that I will throw mine.'

Sentences 13a, 13b, and 14 have essentially the same semantic content. In 13a, however, the top-level pragmatic configuration (Topic-Comment) indicates a "double-focus" contrast whose nodes of origin are the speaker and her sister. In 13b, the top-level configuration shows "single-focus" contrast, pointing out the young man at which the speaker was going to throw her flower. The configuration which is on the top level of structuring in one is shown at an embedded level in the other. In this way, both kinds of contrast are coded in both sentences. Example 14 is similar in its pragmatic configuration to 13b, though simpler.

4.2 Hixkaryana hearsay particle

In Hixkaryana, a Carib language of northern Brazil, the positioning of the hearsay particle is similar to that of the future marker in Mbya

Guarani (Sect. 4.1), in the sense that two types of positioning are involved: one type, in the verb phrase, is grammatically determined and gives no evidence of pragmatic motivation; the second type, as a spacer to indicate a boundary in the pragmatic configuration of an utterance, is fully motivated pragmatically. In this way, pragmatic motivation of sentence positioning operates on a part-time basis. But whereas the positioning of the Mbya future marker as a spacer is not describable by a grammatical rule, the positioning of the Hixkaryana hearsay particle as a spacer can, in fact, be thus described: it occurs in clause-second position.

In Hixkaryana there are five evidential enclitics (called "verification particles" in Derbyshire 1985), each of which has both of the two types of positioning described above: **tɨ** 'hearsay', **na** 'uncertainty', **mpɨnɨ** 'certainty / prediction / warning', **we** 'opinion / recollection / counteraffirmation', and **mpe** 'positive doubt, scepticism'. "The basic order of constituents is OVS, with adjuncts normally following the subject. There is an optional rule which moves the subject or an adjunct to the clause-initial position for purposes of emphasis" (74f). When an evidential occurs, "it is often placed in the verb phrase, which may or may not be the initial phrase of the sentence; however, when a subject or adjunct phrase is fronted for emphasis, the verification particle is usually placed in that initial phrase" (129). In 15a, **tɨ** 'hearsay' (along with **ha** 'intensifier') occurs in the verb phrase; in 15b, it occurs with a locational adjunct which is "fronted for emphasis":

(15) a. mana yonahyatxkon hatɨ, ohoryen heno komo
manna they-ate-it HSY your-ancestor dead COLL
'Your ancestors ate manna.'

b. Emphasis ----Remainder of utterance---
ito tɨ nehxakon ha kamara yohɨ
there HSY he-was INTENSFR jaguar chief-of
'The jaguar chief was there.' (147)

In 16 it is a grammatical subject which is "fronted for emphasis" and followed by **tɨ**. The context is provided in Derbyshire 1986:

(16) Context: The sentence preceding indicated that the sloth was not in the village where he might have been expected to be. That sentence, in which the sloth is referred to by a pronoun, together with the following one, reintroduces the sloth into the discourse:

Emphasis Remainder of utterance
xofrye mah tɨ ehxera n-ehxakonɨ
sloth CTEXP HSY be.NEG 3S-be+DP
'The sloth was not there.'

In discussing 16, Derbyshire states that "the clause-initial, preverbal subject noun phrase [**xofrye** 'sloth'] is a grammatical device that correlates with the semantic factors to indicate that there is some sort of thematic break at this point" (1986:250).

Pragmatic motivation of second positioning is not surprising, since sentence-initial position is often a highly significant one for discourse-pragmatic reasons: the fronting of informative Topics and Focus expressions, for example, is highly universal cross-linguistically (see papers in Givon 1983; Dayley 1985). As Steele puts it, "the tendency of modals, specifically, to sentential second position is a function of the importance of first position" (1975:238).

The Hixkaryana hearsay particle **tɨ**, then, has two sentence positions, both of which can be described separately by a grammatical rule. The overall positioning of **tɨ** is only partially grammaticalized, however, since the choice between those two positions is made on pragmatic rather than grammatical grounds: **tɨ** is pulled out of the verb phrase and placed after a fronted element when it is desired there as a spacer, to highlight the constituent division in a marked pragmatic configuration.

4.3 Hixkaryana afterthought elements

Basic word order in Hixkaryana has been featured in published studies for over ten years (see Derbyshire 1981, 1985, 1986, and references listed therein). Hixkaryana has "obligatory person-marking verb prefixes which agree with subject and object", a "consequent scarcity of noun phrase subjects and direct objects in most discourse contexts", and "OVS as its basic order in clauses that have subject and object noun phrases" (1986:280). When adjuncts occur, they normally follow the subject (1985:74). Breaking the OVS order into its subparts, Derbyshire states that OV is a very rigid order and VS is rigid to a fair degree, in the sense that "OV and VS can, and do, occur in all kinds of discourse-pragmatic environments" (1986:282).

In a comparative study involving other Carib languages, Derbyshire (1981) shows how the present OVS order in Hixkaryana can be seen as the result of drift from earlier SOV; the major motivation for the drift seems to be the grammaticalization of afterthought patterns (Hyman 1975). "Such patterns occur in probably all languages, without necessarily becoming grammaticalized. They occur where a speaker adds something after he has completed the main predication for the purpose of clarification, disambiguation or emphasis, or simply to supply information he had forgotten to include in the main part of the sentence. The afterthought element is phonologically dislocated from the main predication by pause and a new intonation contour, and this signals a highly marked order of sentence constituents. After a period of time, however, the afterthought element may become grammaticalized; that is,

it may become part of the intonation contour of the main predication, without pause, and with a new, relatively unmarked, ordering of syntactic elements" (Derbyshire 1981:216).

In Hixkaryana, the grammaticalization of afterthought elements is "fairly well developed" (1981:218), in the sense that "the normal position of a subject or adjunct phrase is to the right of the verb". There remains, however, "an option for the speaker to make them a part of the main intonation pattern of the clause or to dislocate them" (1985:78). Further, the phonological criteria are often unclear as to whether there is actual right dislocation or not (1986:240). In 17a there is a right-dislocated subject **xofrye** 'sloth', and in 17b non-dislocated final subject **toto heno komo** 'the (former) people'; in 18a there is a right-dislocated adjunct **atunano wya** 'by fever', and in 18b a non-dislocated final adjunct **rowya** 'to me':

(17) a. noseryehokekonɨ, xofrye
he.was.upset sloth
'The sloth was upset' (1986:245)

b. namryekyatxkonɨ toto heno komo
they.went.hunting person dead COLL
'The people used to go hunting' (1985:30).

(18) a. ekeh me wehxaha, atunano wya
sick.one DENOMLZR I.am fever by
'I am sick with fever' (1985:34).

b. yawaka yɨmyako bɨryekomo rowya
axe he.gave.it boy to.me
'The boy gave the axe to me' (1985:35).

In Hixkaryana, then, the pragmatic phenomenon of afterthought elements is grammaticalized sufficiently to establish OVS as the basic word order when full noun phrases are present, but when desired, the speaker can still present such elements as being less than fully integrated into the grammatical structure of the clause.

5 Pragmatic phenomena with more or less full grammaticalization

Pragmatic motivation can exist alongside partial grammaticalization (Sect. 4), or even alongside full grammaticalization. In the latter case, a grammatical rule for the phenomenon is adequate for formal description. What is the point, then, in considering the possibility of pragmatic motivation? There would be no point if one's goal does not go beyond formal description. If, however, it includes giving a description which models "the linguistic intuition of the native speaker" (Chomsky 1964:28) in regard to function as well as form, then all motivation is relevant which can be shown to be operative.

But if a formal rule works, what evidence could there be for system-external motivation? Two types of explanation might be possible. (a) One might be through analogy with phenomena in which the motivation must be recognized on empirical grounds, phenomena such as are described in Sects. 3 and 4. For example, since the pragmatic notion of spacer seems to be necessary to explain the positioning of certain particles in Mbya and Hixkaryana, it might be of interest to consider whether, in another language, certain particles whose positioning is describable in purely grammatical terms (such as second position, Sect. 5.2) could have a certain amount of pragmatic motivation as spacers (cf. Dooley 1987b). (b) A second line of evidence should take into account whether the motivation in question is consistently plausible for all occurrences of the phenomenon. If the argument from motivation shows "leaks", then the purported explanation is not observationally adequate, although pragmatic motivation may have been a factor diachronically. For a fuller discussion of "validating pragmatic explanations," see Nunberg 1981.

In this section, I present two phenomena which are fully grammaticalized. For the first of these, English WH-clefts, it is possible to furnish both of the above lines of evidence, so that full pragmatic motivation can be established; for the second, Dutch auxiliary placement, attempts fall short of that goal.

5.1 English WH-clefts

This section will illustrate that English pseudo- or WH-cleft constructions are fully motivated pragmatically as well as being fully describable grammatically. In this, I follow the analysis of Prince 1978.

Consider the following sentences:

(19) a. Carol's mom made a cake.

b. What Carol's mom made was a cake.

19a is the unclefted form and 19b a corresponding WH-cleft. (In the interests of simplifying the discussion, I am not considering here (it-)clefts such as **It was a cake that Carol's mom made.**) The two sentences in 19 have the same semantic content, but differ along pragmatic lines. The major difference is that 19b conveys a presupposition such as

(20) Carol's mom made something,

whereas 19a conveys no such presupposition. We may speak of 19b as containing a WH-constituent (**what Carol's mom made**) and a predicate (**was a cake**). In a WH-cleft, the presupposition is conveyed by the WH-constituent; such a presupposition, according to Prince, is presented

as GIVEN, assumed to be in the hearer's consciousness at the time of the utterance (903). Consider 21:

(21) a. "Nikki Caine, 19, doesn't want to be a movie star. WHAT SHE HOPES TO DO IS BE A STAR ON THE HORSE-SHOW CIRCUIT." (**Today**, 10/10/76, p. 44, cited in Prince, 887)

b. ...SHE HOPES TO BE A STAR ON THE HORSE-SHOW CIRCUIT.

Both 21a and the more neutral 21b are acceptable in this context. But consider a different context:

(22) a.? Guess who I ran into today — Nikki Caine! WHAT SHE HOPES TO DO IS BE A STAR ON THE HORSE-SHOW CIRCUIT.

b. ...SHE HOPES TO BE A STAR ON THE HORSE-SHOW CIRCUIT.

The acceptability of the cleft construction in 22a sharply diminishes when the context does not make obvious why the presupposition should be presented as given.

Prince discusses subtle contextual conditioning of WH-clefts which we will not go into. The point here is simply that WH-clefting exemplifies a phenomenon which, although consistently motivated, can be formally described as a grammatical phenomenon, without recourse to motivation of any kind. It provides a rather straightforward illustration of Chomsky's (1980) distinction between "grammatical competence" and "pragmatic competence", "restricting the first to the knowledge of form and meaning and the second to knowledge of conditions and manner of appropriate use" (p. 224).

5.2 Dutch auxiliary placement

This final example differs from the preceding ones in that pragmatic motivation cannot be consistently maintained. Examples up until the present have been moving towards the grammaticalized end of the scale for pragmatic phenomena; the following is just off that end of the scale, since the pragmatic motivation it shows is partial and incomplete.

The verb phrase in Dutch often contains more than one word: **heeft geslagen** '(has) hit', **zijn vertrokken** '(have) left', **deed dicht** 'closed (did close)', etc. The first word in a VP is always a finite verb, and when the VP has more than one word, the first is an auxiliary verb. In considering the placement of the auxiliary verb in the independent declarative clause, I follow Kooij 1978.

According to Kooij, independent declarative clauses have 23 as an unmarked order of constituents:

(23) S V1 ... O ... Compl V2 (...)

Here V1 is the first word in the VP, V2 is whatever may remain of the VP, and "Compl is a cover symbol for separable prefixes, directional adverbs, and predicative adjectives." Given 23 as an unmarked ordering, "the ordering of constituents in the Dutch sentence is rather free" (Kooij, 32).

A primary means of expressing this freedom is through fronting, that is, placing a nonsubject constituent in first position. When this is done, the subject moves to a position to the right of V1 (Kooij, 33). Consider the sentences of 24 (Kooij, 30):

(24) a. Ik heb dat verhaal gisteren aan Wim verteld
I have that story yesterday to Bill told
'I told Bill that story yesterday.'

b. Dat verhaal heb ik gisteren aan Wim verteld
that story have I yesterday to Bill told

c. Aan Wim heb ik gisteren dat verhaal verteld
to Bill have I yesterday that story told

d. Gisteren heb ik dat verhaal aa Wim verteld
yesterday have I that story to Bill told

24a has unmarked ordering. In 24b, the direct object is fronted; in 24c, it is an indirect object; and in 24d, a temporal adverb.

According to Kooij, there are three distinct situations in which a constituent can be fronted. Each is described in pragmatic terms: (a) when the constituent is an anaphoric, background expression; (b) when it is a sentential-scope expression of setting or modality; or (c) when it is a new or contrastive focus expression. Type (a) fronting can be illustrated by sentences 24b and 24c, in a context in which both the subject **ik** 'I' and the direct object **dat verhaal** 'that story' are backgrounded as given information, not presented as highly informative. Type (b) fronting can be illustrated by 24d. Another illustration of type (b) is the following (Kooij, 45):

(25) Gelukking deed Jan op tijd de deur dicht
fortunately did John in time the door close
'Fortunately, John closed the door in time.'

Type (c) fronting can be illustrated by the following sentences (Kooij, 36):

(26) Hard heeft-ie niet gewerkt
hard has-he not worked
'He did not exactly work hard.'

(27) Een auto hebben ze gisteren aam Wim gegeven!
a car have they yesteday to Bill given
'It's a car they gave Bill yesterday!'

As the added accent marks indicate, fronted constituents of this type receive phonological prominence. Sentence 26 would be appropriate in a contrastive situation that could be paraphrased by 'You may claim that the boy worked hard, but I claim that he did not' (Kooij, 36).

Since fronting is common in Dutch (much more so than in English, Kooij, 40), and since it is also common for the auxiliary verb to follow the fronted element, it would not be unreasonable to ask if the placement of the auxiliary had pragmatic motivation. The question has more point when one compares the auxiliary element with Mbya and Hixkaryana spacer particles (Sects. 3.1, 4.1, and 4.2). Like spacer particles, Dutch (and many other) auxiliaries are typically short, unstressed elements with largely sentence-scope semantics (tense, aspect, and mood). Recall, as well, that spacer particles function as spacers just by means of these properties, along with one other: their positioning between components of a plausible pragmatic configuration (Sect. 3.1). In regard to type (c) fronting, Kooij states that the fronted focus elements "are, as it were, cut off from the rest of the sentence by intonation, and... by the finite verb" (p. 37). Let us, then, entertain the hypothesis that the placement of Dutch auxiliary verbs is pragmatically motivated, and that they function as spacers.

Notice first of all that in order to dismiss this hypothesis, it is not sufficient merely to make the point that the placement of Dutch auxiliaries is grammatically controlled, fully describable by grammatical rules. As we saw in Sect. 5.1, it is possible for full pragmatic motivation to coexist with grammatical control. Rather, we must show that the putative pragmatic motivation is not consistently the case; the pragmatic explanation leaks.

In the present case, that is not difficult to do. We simply observe that the auxiliary occurs following subjects or other constituents which cannot be plausibly maintained to be components in a pragmatic configuration. Consider, for example, dummy subjects (Verhagen, 179):

(28) Er sal waarschijnlijk in november met de
there will probably in November with the

afwerking worden begonnen
finishing.touches get started

'Probably the finishing touches will start in November.'

It would be difficult to imagine motivation which would justify a dummy subject as a constituent in a pragmatic configuration.

There are two other considerations which mitigate against considering the auxiliary as a consistent boundary marker between pragmatically significant constituents. First, the second positioning for auxiliaries is but a subcase of a broader rule which says, "The finite verb invariably occupies the second position" (Kooij, 30). That is, a functional explanation for auxiliary placement would need to address the placement of finite verbs in general, and that would present even more difficulties, for example in regard to prototypical properties of spacers. Second, there is some evidence that modal adverbs regularly occur at the boundary between topic and comment. This point is made implicitly by Kooij (p. 35) and explicitly by Verhagen (1986:96ff). Consider sentence 29 (Verhagen, 98):

(29) Marie zal hem waarschijnlijk afwijzen
Marie will him probably reject
'Marie will probably reject him.'

Here the position of the modal adverb **waarschijnlijk** 'probably' indicates that both **Marie** and **hem** are background material, not in the comment (the most informative part of the utterance), which consists only of **afwijzen** 'reject'. If this is so, it would be hard to maintain that the auxiliary **zal** indicates another kind of pragmatically relevant boundary.

We are left, then, with a leaky pragmatic explanation. It might be interesting to investigate whether pragmatic motivation was active diachronically in assigning the auxiliary to second position. It does seem likely that this positioning carries real-time pragmatic motivation in certain situations, such as type (c) fronting referred to above. But as a global explanation, it appears inadequate to say that real-time pragmatics provides motivation for the second positioning of Dutch auxiliaries.

6 Concluding remarks

In this paper I have attempted to show that (a) there are externally motivated -- specifically discourse-pragmatic -- phenomena which are encoded in the syntax and fully controlled by the grammar; (b) there are phenomena of this type which are encoded in the syntax but are only controlled by the grammar in some minimal sense; and (c) there are phenomena of this type which are encoded in the syntax and controlled by the grammar to varying degrees and in different ways. Thus, this study assumes a model of language description in which pragmatics and grammar can, on the one hand, compete for control of syntactic phenomena or, on the other, cooperate, to the extent that pragmatic motivation and grammatical patterning are mutually compatible.

The mere fact that grammaticalization of pragmatic phenomena takes place does not argue for a purely formal approach to syntax any more than it argues for a functional one; grammaticalization itself can often be seen as a particular instance of conventionalization, a phenomenon broadly observable in many aspects of human activity. On the other hand, the fact that there exist system-external explanations for certain syntactic facts cannot be taken to imply that other facts should not be explained internally, at least in the synchronic, real-time sense. Indirectly, then, a study of this kind constitutes an appeal for linguistic description that allows for both functional and formal approaches. Reductionism in either direction -- the positing of one model to the exclusion of the other -- though it might make for exciting polemics, lends itself to a caricature of language.

NOTES

1. I would like to acknowledge helpful suggestions given by a number of people in the preparation of this paper, particularly Albert Bickford, Des Derbyshire, and Gloria Kindell. In this regard I have two regrets: first, that time and geography have not allowed the input of others, and second, that in this paper it has not been possible to address all of the important issues raised by comments I did receive. Shortcomings remaining in the paper are, of course, my own responsibility.

2. This paper makes use of various sources of data, and only light editing was done on the glosses of examples, principally limited to formatting; the original authors' preferences were followed in matters of style and abbreviation. The following abbreviations occur:

ANA	anaphora
CONT	continuative
COLL	collective
CTEXP	contrary to expectation
DAT	dative
DENOMLZR	denominalizer
DP	distant past
DS	different subject
EXA	exactly
FUT	future
HSY	hearsay
IMPER	imperative
INTENSFR	intensifier
NEG	negation
OP	opinion of speaker
PL	plural
Q	question marker
REL	relativizer
SER	serial verb suffix
SG	singular
SS	same subject
1	first person
2	second person
1+2	first person plural, inclusive
3	third person

REFERENCES

Austin, J. L. 1962. How to do things with words. Oxford: Clarendon Press.

Carnap, Rudolf. 1942. Introduction to semantics. Cambridge, MA: Harvard U. Press.

Chomsky, Noam. 1964. Current issues in linguistic theory. The Hague: Mouton.

_____. 1971. Deep structure, surface structure, and semantic interpretation. Semantics, ed. by Danny D. Steinberg and Leon A. Jakobovits, 183-216. Cambridge: Cambridge University Press.

_____. 1975. Reflections on language. New York: Pantheon.

_____. 1980. Rules and representations. New York: Columbia University Press.

_____. 1982. The generative enterprise: a discussion with Riny Huybregts and Henk van Riemsdijk. Dordrect: Foris.

Cole, Peter, ed. 1981. Radical pragmatics. New York: Academic Press.

_____ and Jerry L. Morgan, eds. 1975. Syntax and semantics, vol. 3: speech acts. New York: Academic Press.

Dayley, Jon P. 1985. Why all languages aren't SOV or VOS, or How competing motivations lead to natural inconsistency. Proceeding of the Eleventh Annual Meeting of the Berkeley Linguistics Society, 52-64. Berkeley, CA: Berkeley Ling. Society.

Derbyshire, Desmond C. 1981. A diachronic explanation for the origin of OVS in some Carib languages. Journal of Linguistics 17:209-20.

_____. 1985. Hixkaryana and linguistic typology. Dallas: Summer Institute of Linguistics and Arlington: U. of Texas at Arlington.

_____. 1986. Topic continuity and OVS order in Hixkaryana. Native South American discourse, ed. by Joel Sherzer and Greg Urban, 237-306. Berlin: Mouton de Gruyter.

Dooley, Robert A. 1982. Options in the pragmatic structuring of Guarani sentences. Lg. 58.307-31.

_____. 1983. Grammaticalization: an introduction. Notes on Linguistics 28.4-12.

_____. 1987a. Basic configurations of pragmatic structuring. 1987 Work Papers of the Summer Institute of Linguistics, University of North Dakota session, volume 31, ed. by Desmond C. Derbyshire, 1-27. Grand Forks, ND: SIL.

_____. 1987b. The positioning of non-pronominal clitics and particles in lowland South American languages. Paper presented at the Working Conference on Amazonian Languages, U. of Oregon, Eugene, Oregon, August, 1987.

Fillmore, Charles J. 1981. Pragmatics and the description of discourse. In Cole, pp. 143-66.

Gazdar, Gerald. 1979. Pragmatics: implicature, presupposition, and logical form. New York: Academic Press.

Givon, Talmy. 1979. From discourse to syntax: grammar as a processing strategy. Syntax and semantics, vol 12: discourse and syntax, ed. by Talmy Givon, 81-112. New York: Academic Press.

_____, ed. 1983. Topic continuity in discourse: a quantitative cross-linguistic study. Amsterdam: John Benjamins.

Grice, H. P. 1968. Utterer's meaning, sentence meaning and word meaning. Foundations of Language 4:225-42.

_____. 1975. Logic and conversation. In Cole & Morgan, 41-58.

_____. 1978. Further notes on logic and conversation. Syntax and semantics, vol. 9: pragmatics, ed. by Peter Cole, 113-28. New York: Academic Press.

Haiman, John. 1985. Natural syntax: iconicity and erosion. Cambridge: Cambridge U. Press.

Hyman, Larry M. 1975. On the change from SOV to SVO: evidence from Niger-Congo. Word order and word order change, ed. by Charles N. Li, 113-48. Austin: U. of Texas Press.

_____. 1984. Form and substance in language universals. Explanations for language universals, ed. by Brian Butterworth, Bernard Comrie, and Osten Dahl, 67-85. Berlin: Mouton.

Jensen, Cheryl. 1980. Word order in Olampi. ms, University of Oregon.

Katz, Jerrold J. 1977. Propositional structure and illocutionary force. New York: Crowell.

Kooij, Jan G. 1978. Fronting in Dutch. Studies in fronting, ed. by Frank Jansen, 29-51. Lisse: Peter de Ridder Press.

Leech, Geoffrey N. 1983. Principles of pragmatics. London: Longman.

Morris, Charles. 1938. Foundations of the theory of signs. International Encyclopedia of Unified Science, Vol. 1, No. 2. Chicago: U. of Chicago Press.

_____. 1946. Signs, language and behavior. New York: Prentice-Hall.

Newmeyer, Frederick J. 1983. Grammatical theory: its limits and its possibilities. Chicago: University of Chicago Press.

Nunberg, Geoffrey. 1981. Validating pragmatic explanations. In Cole, 199-222.

Osgood, Charles E. and Thomas A. Sebeok. 1954. Psycholinguistics: a survey of theory and research problems. IJAL, Memoir 10. Baltimore: Waverly Press.

Payne, Doris L. 1987. Meaning and pragmatics of order in selected South American Indian languages. Paper presented at Conference no. 105, Wenner-Gren Foundation for Anthropological Research, The Role of Theory in Language Description, Ocho Rios, Jamaica, Oct. 31-Nov.8, 1987.

Prince, Ellen F. 1978. A comparison of WH-clefts and it-clefts in discourse. Lg. 54.883-906.

Searle, John R. 1969. Speech acts: an essay in the philosophy of language. Cambridge: Cambridge U. Press.

_____. 1975. Indirect speech acts. In Cole & Morgan, 59-82.

_____. 1976. A classification of illocutionary acts. Language in Society 5.1-23.

_____, Ferenc Kiefer, and Manfred Bierwisch, eds. 1980. Speech act theory and pragmatics. Dordrecht: Reidel.

Sperber, Dan and Deirdre Wilson. 1986. Relevance: communication and cognition. Oxford: Basil Blackwell.

Stalnaker, Robert. 1972. Pragmatics. Semantics of natural language, ed. by Donald Davidson and G. Harman, 380-97. Dordrecht: Reidel.

Steele, Susan. 1975. On some factors that affect and effect word order. Word order and word order change, ed. by Charles N. Li, 197-267. Austin: U. of Texas Press.

Thompson, Sandra A. 1978. Modern English from a typological point of view: some implications of the function of wort [sic] order. Linguistische Berichte 54.19-35.

Verhagen, Arie. 1986. Linguistic theory and the function of word order in Dutch: a study on interpretive aspects of the order of adverbials and noun phrases. Dordrecht: ICG Printing.

Verschueren, Jef, ed. 1987. Concluding round table, 1987 International Pragmatics Conference. IPrA Working Document 2. Antwerp: International Pragmatics Association.

ANAPHORIC INDICES AND INALIENABLE POSSESSION IN BRAZILIAN PORTUGUESE

Daniel Everett

In this squib, I want to draw attention to some previously unnoticed but intriguing facts about inalienably possessed NPs (INPs) in Brazilian Portuguese (BP). I argue that these differ from both pronouns and anaphors in that (i) they must always be c-commanded in their minimal Complete Functional Complex (CFC) by an argument with an independent θ-role and (ii) the Binding Theory (BT) applies to IPNPs algorithmically: first the BT attempts to treat them as anaphors; if successful, the BT is satisfied; if not, then it treats them as pronouns[1]. First, let us examine the evidence for (i).

IPNPs are like anaphors in that they may not appear as unaccusative subjects, (1) and (2). Unlike anaphors, however, they may appear as passive subjects, (3):

(1) *Se chegou. (reflexive in unaccusative construction)
'Self arrived.'

(2) *A (sua) perna de (Sérgio) {operou / quiemou / quebrou} (por João).[2]

'The (3rd person's/Sérgio's) leg {operated on / burned / broken} (by John).'

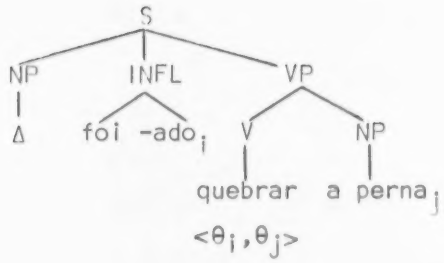
(3) a. A perna (de Sérgio) foi operada/etc. (por João).
'The (Sérgio's) leg was operated on/etc. (by John).'

b. *Se foi operado.
'Self was operated on.'

If IPNPs and reflexives are both anaphors, (1) and (2) are expected since there is no c-commanding antecedent available in unaccusative constructions. But, if both are anaphors, why may inalienably possessed forms (3a), but not reflexives (3b) appear as passive subjects? The answer is that IPNPs and reflexives are not exactly alike: whereas reflexives, as anaphors, require an antecedent (in the appropriate domain, see (7) below), IPNPs merely need to be c-commanded by an argument.

To see how this applies to (1)-(3), recall how the passive differs from unaccusatives. Some researchers, e.g. Baker (1988) and Everett (1986), have argued that the passive morpheme, -ado in BP, is an argument, generated under INFL, receiving the external θ-role of the verb, as in (4):

(4)



Given (4) and the stipulation that IPNPs in BP must be c-commanded in their minimal CFC by an argument with an independent θ-role, we account for the contrasts in (1)-(3). The CFC of an unaccusative verb will not contain a c-commanding NP for its subject but the CFC of a passive will. Note that the reference to CFC is crucial:

(5) *João disse [que a (sua) perna (de Sérgio) quebrou].
'John said that the (his/Sérgio's) leg broke.'

In (5) João c-commands perna 'leg', but the example is still ungrammatical since João occurs outside the CFC of perna. As shown, the appearance of an optional possessor is not enough to satisfy the CFC requirement, whether the possessor is pre- or postnominal. This results from the fact that the IPNP as a whole must be c-commanded. Although a possessor c-commands its head, it does not c-command the IPNP which dominates it, failing to save the structure. How can this requirement be formalized? I submit that a means is already available, namely, anaphoric indices.

In Chomsky (1980), two types of nominal indices were distinguished: the referential index, which indicated the reference of the NP bearing it, and the anaphoric index, the set of referential indices of those NPs c-commanding the NP in question. For example:

(6) John$_i$ said that Bill$_{j\{i\}}$ hit him$_{k\{i,j\}}$.

Since no NP c-commands John, it bears no anaphoric index. Bill's anaphoric index contains the referential index of John, which c-commands it. Likewise, him contains in its anaphoric index the referential indices of John and Bill, both of which c-command it. Chomsky (1980) proposed to account for the binding properties of sentences via the interaction between referential and anaphoric indices.

After Chomsky (1981), however, anaphoric indices were abandoned as superfluous, their effects subsumed under the Binding Conditions. Nevertheless, the data here suggest that something like the anaphoric index must be retained. An account of reflexive facts in BP must express the fact that IPNPs, although not required to be bound, must be c-commanded by a nominal argument, exactly what is expressed by

anaphoric indices. The intuition is that an IPNP will appear to be disembodied (and such structures are no longer interpreted by native speakers as possession structures but as weird or ungrammatical) unless a potential possessor is available in the form of an independent argument. This can be formalized as in (7):

(7) For α , an inalienably possessed noun, there must be a β , Domain of Possession, such that α assigns θposs to γ , NP, in β , where:
β is a Domain of Possession iff:
a. β is the minimal CFC containing α and
b. the anaphoric index of α is nonnull in β .[3]

We have now established and formalized (i) above, arguing for the retention of anaphoric indices. Now, let us examine claim (ii), that the BT applies algorithmically to IPNPs.

The notions of Binding Theory and Binding Domain (BD) are given by Chomsky (1986:171-172):

(8) "a. The indexing I and the pair (α , β) are compatible with respect to the binding theory (BT) if α satisfies the binding theory in the local domain β under the indexing I:
I is BT compatible with (α , β) if:
(A) α is an anaphor and is bound in β under I
(B) α is a pronominal and is free in β under I[4]

b. The licensing condition for a category governed by a lexical category in the expression E with indexing I:
For some β such that α is an anaphor or pronominal and β is the least functional complex (CFC) containing γ for which there is an indexing J BT-compatible with (α , β), I is BT-compatible with (α , β)."

Previous studies, such as Burzio (1986,265ff), have argued that inalienably possessed NPs are like reflexives in that their antecedent (i.e., possessor) must occur within the minimal S in which the pronoun occurs:

(9) *Maria$_i$ ha detto [che Giovanni alzasse la mano$_i$].
'Maria$_i$ said that Giovanni should raise her hand.'

This example is bad, according to Burzio, because the possessor is outside the BD of the IPNP. The same generalization holds true for Portuguese[5].

(10) Maria$_i$ disse [que João deve levantar a mão$_i$].
'Maria$_i$ said that Giovanni should raise her$_i$ hand.'

(11) a. Maria$_i$ disse[que a mão$_i$ foi levantada(por João)].
'Maria$_i$ said that{her$_i$ / someone's} hand was raised (by João).'

b. ??/*Maria disse [que a mão levantou.]
'Maria said the hand raised.'

In (11a), but not in (10), the antecedent of the possessive pronoun may appear in the matrix clause. While in (11b) only the disembodied, non IPNP reading is available, (11a) admits several interpretations. Mão may be disembodied, a non IPNP reading, or it may belong to an arbitrary possessor, including Maria. It is the possessor reading that we are interested in. This reading is grammatical because mão occurs in a DP and although it has no antecedent in its BD, it is nonetheless grammatical since it may alternatively be treated as a pronominal (i.e. its possessor position may be, see note 5). Given these facts, we can define the conditions on inalienable possession in BP as in (12):

(12) a. For any α , α an IPNP, there must be a DP;
b. If BC-A can apply to α , it must apply.
c. If BC-A cannot apply to α , apply BC-B.

Note that (11) shows that it is not the passive morpheme itself which binds the possessed noun. The mão belongs to Maria but Maria is not the agent of levantar, which would have to be the case if the real binder of the possessed noun were the passive morpheme. In other words, mão is subject to BC-A IN (10) and BC-B in (11), just as predicted by (12). There is no 'implicit' binder.

Thus, the facts of inalienably possessed nouns in BP can be expressed simply - but only if anaphoric indices are retained in the theory.

NOTES

1. Chomsky (1986:169) defines Complete Functional Complex (CFC) as the domain in which "...all grammatical functions compatible with its head are realized in it - the complements necessarily, by the projection principle, and the subject, which is optional unless required to license a predicate, by definition." By IPNP, I refer to human body parts, although this probably does not exhaust the class. I will not consider in detail the role of pre- or postnominal possessors in IPNPs, but see note 5.

2. As discussed in Everett (in preparation), operar and the other verbs listed here may appear in unaccusative-like constructions in BP.

3. Verbs of sensation, e.g. esquentar 'to warm up'; sangrar 'to bleed'; and cocar 'to itch', violate this. In Everett (in preparation),

I argue that this is because their lexical structure renders them inherent DPs.

4. In Everett (in preparation), I argue that IPNPs always assign a θ_{poss} (possessor) role in a DP, as per (7). Consequently, IPNPs will always be CFCs and BDs. Thus, pronominal possessors of IPNPs may be bound, as in John$_i$ broke his$_i$ arm, without violating (8). This means that IPNPs will always have an overt or ec possessor, although I take no stand here as to what kind of ec that might be. I will refer to (A) of this definition as BC-A and (B) as BC-B.

5. Actually, since these binding facts hold only for IPNPs without an overt possessor, the real generalization must ultimately be stated in terms of the IPNP possessor rather than its head. When an overt possessor is present, the IPNP is subject to BC-B if it is pronominal and is treated as a name if the possessor is an R-expression (Chomsky 1982) cf. also note 4). I avoid this issue here though.

REFERENCES

Baker, Mark. 1988. Incorporation: A theory of grammatical function changing. University of Chicago Press.

Burzio, Luigi. 1986. Italian syntax: A government-binding approach. Dordrecht: D. Reidel.

Chomsky, Noam. 1980. On binding. LI. 11:1-46.

Chomsky, Noam. 1981. Lectures on government and binding. Dordrecht: Foris.

Chomsky, Noam. 1982. Some concepts and consequences of the theory of government and binding. Cambridge, MA: MIT Press.

Chomsky, Noam. 1986. Knowledge of language: its nature, use, and origins. New York: Praeger.

Everett, Daniel. 1986. 'Pirahã clitic doubling and the parametrization of nominal clitics. MIT working papers in linguistics, ed. by N. Fukui, T. Rapoport, and E. Sagey. Vol. 8:85-127.

Everett, Daniel. in preparation. 'Some theoretical consequences of transitive verb classes in Brazilian Portuguese. ms, University of Pittsburgh.

Williams, Edwin. 1980. 'Predication'. LI. 11:203-238.

CLITICS, CASE AND WORD ORDER IN YAGUA

Daniel L. Everett

INTRODUCTION
1 Overview of Yagua word order
2 Clitics
  2.1 Set I clitics
  2.2 Set II clitics
3 Parametrization of Yagua clitics
4 Case and word order
  4.1 Introduction
  4.2 Case assignment
  4.3 Multi-AGR languages
  4.4 Case assignment and doubling in VP
    4.4.1 Direct objects
    4.4.2 Indirect objects
  4.5 Definiteness
  4.6 Summary
  4.7 Unaccusative subjects
  4.8 Case, subjects, possessors and obliques
    4.8.1 Subjects
    4.8.2 Possessor doubling
    4.8.3 Postpositional phrases
    4.8.4 Conclusion
5 Reflexives
6 Underlying versus basic word order
  6.1 Underlying word order
  6.2 Basic word order
7 Some diachronically oriented speculation
8 Implications for lexical theories
9 Conclusion
NOTES
REFERENCES

Introduction

In the Peruvian language, Yagua, whenever a nominal clitic appears in NP, S, or PP, a tautophrasal, coreferent NP may no longer occur in its normal phrase initial position but must instead be placed after its phrasal nucleus, as in (1)-(3):

(1) a. Pauro púúchu Anita. SUBJECT
Paul carry Anita
'Paul carries Anita.'

b. Sa_i-púúchu $Pauro_i$-nii_j $Anita_j$.
3SGI-carry Paul-3SGII^j Anita^j
'Paul carries Anita'.

c. *$Pauro_i$ sa_i-púúchu(-nii_j) $Anita_j$.

(2) a. Alchico rooriy. GENITIVE NP
Alchico house
'Alchico's house.'

b. Sa_i-rooriy $Alchico_i$.
'Alchico's house.'^{i}

c. *$Alchico_i$ sa_i-rooriy.

(3) a. Nurutu viimu. POSTPOSITIONAL NP
alligator inside
'Inside the/an alligator.'

b. Sa_i-viimu $nurutu_i$.
'Inside the/an alligator.'

c. *$Nurutu_i$ sa_i-viimu.

Clitic doubling does not alter the postverbal order of objects, as seen in (1b). Clitics doubling objects differ from the clitics in (1)-(3) in that they are suffixed to their host, do not produce word order changes, and show a lesser degree of attachment to their host than clitics doubling possessors, subjects, or postpositional objects. Further, doubled objects, but not other doubled NPs, are interpreted as definite. Clitics which double objects in Yagua are called Set II clitics. All other clitics are referred to as Set I clitics.

In this paper, I argue that the relationship between Set I clitics and Set II clitics, and the effect of these clitics on word order, present some intriguing problems for theories of syntax, especially as seen in the phenomenon of clitic doubling, the co-occurrence of tautophrasal, coreferent clitic NP pairs associated with a single semantic or thematic (θ)-role. I argue that a careful study of Yagua clitic doubling makes at least the following contributions to linguistic theory:

1. The crosslinguistic parametrization of nominal clitics proposed in Everett 1986, 1987 receives significant support. Simply put, clitics may vary crosslinguistically in whether they require a θ-role or morphosyntactic case (henceforth Case).

2. By establishing that pragmatically unmarked word order (VSO) and syntactically underlying word order (SVO) in Yagua are coexistent, mutually compatible components of Yagua grammar, the concept of basic word order, assumed implicitly by the majority of typological studies, is shown to be of little use, unless defined more precisely in terms of the notions just mentioned. Incidentally, by demonstrating that SVO is the underlying order, then, contrary to the claims of Payne (1986), Yagua violates no typological universals.

3. Yagua Set I clitics are sensitive to otherwise word internal morphophonological processes, such as Vowel Harmony, which do not affect Set II clitics. However, as we will see, there are strong arguments for attaching all clitics in the syntax, i.e., postlexically. This represents a serious problem for models which account for word internal phonology or inflection in the lexicon, e.g. Lexical Phonology (LP. Mohanon 1986, Kiparsky 1985, Pulleyblank 1986) and Lexical-Functional Grammar (LFG. Bresnan 1982). It supports instead the attachment of at least some morphemes in the syntax.

The discussion is organized as follows. First, we survey the basic facts of Yagua clitic doubling and its interaction with word order. Section two presents additional facts about Yagua clitics which must be accounted for. In section three, it is shown that Yagua clitics are nonarguments. In section four, I propose an analysis of the facts in terms of Government and Binding Theory (GB). Section five adduces independent evidence for this analysis from some otherwise puzzling facts of Yagua reflexivization. The implication of this analysis for syntactic typology, and theories such as LP and LFG are discussed in the final sections. Results and major features of this study are summarized in the conclusion.

1 Overview of Yagua word order

D. Payne (1985) notes that it is not immediately apparent whether 'basic word order' is SVO or VSO. Both orders are natural and occur frequently. However, due to the fact that SVO order is more salient in discourse (i.e., it has special discourse functions such as introducing new participants), she concludes that VSO, the less salient order, is 'pragmatically unmarked' and hence the basic word order. Interestingly, however, as we have already seen in (1)-(3) above, there are other differences between these orders which cannot be explained exclusively in terms of pragmatic, discourse related factors.

Payne 1986 observes that full NP arguments are rare in Yagua discourse. When they are absent, however, a corresponding clitic is obligatory:

(4) a. Sa-juuy.
3SGI-fall
'He/she falls.'

b. *Juuy.

(5) a. Sa-siiy.
3SGI-run
'He/she runs.'

b. *Siiy.

The clitic may be omitted, however, when a full NP complement is present:

(6) Anita juuy.
'Anita runs.'

(7) Davi siiy.
'David runs.'

When a NP complement occurs without a coreferent clitic, the required order is preverbal for subject, prenominal for genitive, and prepositional for adpositional objects (henceforth obliques), as in (1)-(3) above. This is further supported by the ungrammaticality of examples like (8)-(10):

(8) *Juuy Anita.
'Anita runs.'

(9) *Rooriy Alchico.
'Alchico's house.'

(10) *Viimu nurutu.
'Inside the/an alligator.'

In light of examples (1)-(10), the analysis of phrasal constituent orders cannot be relegated exclusively to pragmatics. Sharp, constant grammaticality contrasts which are not significantly altered by context are unlikely to be pragmatic in nature. At a very minimum, we must investigate the plausibility of a syntactic account for the above restrictions. Examples like (1) above and (11) below illustrate the fact that objects always follow the verb (except in Left-Dislocation structures, which are orthogonal to our present concerns. See D. Payne 1985:28ff for details).

(11) a. Rospita suuta Anita.
'Rospita washes Anita.'

b. Sa_i-suuta $Rospita_i$-nii_j $Anita_j$.

c. {*Anita Rospita / *Rospita Anita} suuta(-nii).

d. *$Anita_j$ sa_i-suuta $Rospita_i$(-nii_j).

Subjects, genitives, and obliques (prepositional objects) thus pattern similarly. Only these grammatical functions may be doubled by Set I clitics. As we will see directly, Set II clitics, which double direct objects, are like Set I clitics in that their host must immediately precede their double.[1] Therefore, a grammar of Yagua must contain and, ideally, explain (12):

(12) The host of a clitic must immediately precede the clitic's double.

Any grammar of Yagua must express the fact that the direct object always appears to the right of its head (V), while genitives and obliques always precede (unless doubled) their heads, N and P, respectively:

(13) a. V^O - Object
b. Genitive - N^O(~ clitic-N^O Genitive)
c. Oblique - P^O(~ clitic-P^O Oblique)

That is, verbs always precede their objects while nonverbs (N and P) follow nondoubled complements and precede doubled complements. An attempt to explain (12) and (13) in terms of independently necessary grammatical principles must wait until section four, after all the relevant data has been presented. Let us turn now to consider some more facts about Yagua clitics.

2 Clitics

2.1 Set I clitics

As we have seen, Set I clitics may be prefixed to V, N, and P. They may double subjects, genitives, and obliques. Moreover, both D. Payne (1985) and T. Payne (1983; 1985) note that Set I clitics are 'more closely attached' to their host than Set II clitics. This observation is based on the fact that Set I clitics undergo certain otherwise word internal phonological processes with their host, such as Vowel Harmony, which Set II clitics fail to undergo. D. Payne (1987:7) says that when the 3SG. Set I clitic sa is attached to an /h/-initial root whose first vowel is any other than /o/ or /e/, the vowel of the clitic, /a/, is changed to '...a vowel of the quality of the first root vowel...' If the

first root vowel is /o/ or /e/, no change occurs. (/h/ = orthographical j; /o/ = orthographical u):

(14) a. Sa-rupiiy (no change) 'He/she walks.'
b. Sa-jimiiy→Simiiy 'He/she eats.'
c. Sa-juunay→Suunaay 'He/she cries.'

These examples further show that Set I clitics attach only to the head of the phrase of which their double is the complement. One way of expressing this in GB terms is to say that the host of a Set I clitic must assign a θ-role to the double. To summarize the facts concerning Set I clitics which must be accounted for: (i) they are prefixed to their host; (ii) their host assigns a θ-role to their double; (iii) their host must immediately precede their double; (iv) they undergo word internal phonological processes with their host; (v) when a Set I clitic is present, and only then, the genitive, oblique, and subject NPs must follow N, P, and V, respectively. Otherwise, they precede these elements.

Set I clitics are listed in Table One:[2]

SET 1 CLITICS

	1	1+2	2	3
Singular	ray-	--	jiy-	sa-
Dual	naay-	vuuy-	saada-	naada-
Plural	nuuy-	vuuy-	jirey-	riy-

TABLE ONE

2.2 Set II clitics

Set II clitics differ from Set I clitics in a number of ways. First, they may only double VP internal NPs. Second, their host need not assign a θ-role to their double. They may thus appear on any word immediately preceding their double. Doubles of Set II clitics are interpreted as more definite, less generic than nondoubled direct objects. Set II clitics are enclitics whereas Set I clitics are proclitics. Set II clitics, unlike Set I clitics, fail to undergo word internal phonological processes with the host. The following examples illustrate these properties:

(15) Sa$_i$-puuchu(*-nîî) Pauro$_i$ rooriy(*-nîî)
3SGI-carry Paul house

viimu-nîî$_j$ Anita$_j$.
into-3SGII Anita
'Paul carries Anita into the house.'

In (15), we see that the clitic must immediately precede its double, as shown by the ungrammaticality of placing the clitic anywhere else. This is an adjacency restriction. The clitic may be attached to any of the potential hosts in (15), so long as it immediately precedes Anita:

(16) a. Sa-puuchu(*-nîî) Pauro-nîî Anita.

b. Sa-puuchu-nîî Anita.

The definiteness of doubled objects is shown in (17) and (18):

(17) a. Sa-jatu buyaa.
3SGI-drink manioc:beer
'He/she drinks manioc beer.'

b. Sa-jatu-ra$_i$ buyaa$_i$.

3SGINII

'He/she drinks the manioc beer.'

This example shows that doubling by a Set II clitic produces a definite reading. Example (18) further illustrates the definiteness of doubled objects, seen in the impossibility of doubling a partitive NP:

(18) a. Jir-rimiy-maa ray-rááva-ta.
2SGI-spill-perfective 1SGI-poison-partitive
'You spilled part of my poison.'

b. *Jir-rimiy-maa-ra$_i$ ray-rááva$_i$-ta.

3SGINII

Partitive case, marked by -ta, is semantically indefinite and thus cannot be combined with a marker for definiteness (the Set II clitic), since this results in a contradictory specification for the same feature. The facts of Set II clitics and their contrast with Set I clitics is explained in the next section. Set II clitics are summarized in Table Two:

SET II CLITICS

	1	1+2	2	3
Singular	-ray	--	-jiy	-nii 'animate' -ra 'inanimate'
Dual	naay	-vuuy	saada	naada
Plural	nuuy	-vuuy	jirey	-riy

TABLE TWO

Before proposing an analysis of these facts, let us summarize our observations to this point, in order to understand as clearly as possible what we must account for:

1. VSO order occurs if and only if the subject is doubled.
2. SVO order is required for nondoubled, subject NPs.
3. Genitive-N^{o} order is obligatory when the genitive NP is not doubled.
4. When the genitive is doubled, the required order is clitic-N^{o} NP.
5. Oblique-P^{o} order is obligatory when the oblique NP is not doubled.
6. When the oblique is doubled, the required order is clitic-P^{o} NP.
7. For all clitics, the host must immediately precede the double.
8. The host of a Set I clitic must θ-mark the clitic's double.
9. The host of a Set II clitic need not θ-mark the double.
10. Set I clitics are proclitics.
11. Set II clitics are enclitics.
12. Set I clitics are affix-like in that they undergo word internal phonological processes with the root. They manifest a closer degree of attachment to their host than Set II clitics.
13. NPs doubled by Set II clitics are less generic, more definite in interpretation.
14. A clitic is obligatory (for both sets of clitics) when no double is present, optional otherwise.
15. As per Payne 1983, the requirement that Set II clitics must be right-adjacent to their double indicates a kind of syntactic constituency between the clitic and its double to its right. On the other hand, Set II clitics form a phonological constituent with their host, on their left. The result is what Payne 1983 terms 'wrong way' cliticization. Ideally, this too should follow from independent principles of the analysis.

An additional fact, which has not yet been commented on but which may be seen easily in examples like (18a), is that tense-aspect markers are suffixed whereas subject agreement is prefixed to its host. This is a common fact about Amazonian languages, most of which derive from an

SOV word order historically (Derbyshire 1986). In favor of the GB approach proposed below is that this property of Yagua is derived from independently necessary principles of Case assignment. Let us turn now to consider some specific proposals.

3 Parametrization of Yagua clitics

Although many authors have considered nominal clitics to be nonarguments universally (see especially Borer 1984 and the papers in Borer 1986), at least two researchers have argued that in fact clitics vary crosslinguistically with respect to argumenthood, Everett (1986, 1987) and Aoun (1981). If clitics do vary crosslinguistically, we must determine whether Yagua clitics are arguments or not. The evidence suggests that they are not. To see why, consider once more clitic-doubling examples, as illustrated in (19) and (20) and previous:

(19) Sa$_i$-dííy Tomása$_j$-ra$_j$ ravichú$_j$-iva.
3SGI-see Tom-3SGINII rock-dat
'Tom saw the rock.'

(20) Sa$_i$-dáátya-nu Pauro$_j$-níí$_j$
3SGI-know-transitivizer Paul-3SGII

Antonio$_j$-ra$_k$ niquee-jada$_k$
Antonio-3SGII talk-infinitive
'Paul teaches Antonio the word.'

The independent referring expressions in these examples (all common or proper nouns) are either the arguments of their clauses, receiving the θ-roles assigned by their phrasal heads or they are appositional, intraphrasal topics in nonargumental (A')-positions[3]. These are the only possibilities. Whatever the role of the doubled NP, it will severely restrict the classification of the clitic. This is so because a single head may not have two arguments associated with the same θ-role. This restriction is known as the θ-Criterion (Chomsky 1982:6):

(21) θ-Criterion: Each term of LF that requires a θ-role (each argument) [emphasis in original, DLE] is assigned a θ-role uniquely. Each θ-role determined by lexical properties of a head is uniquely assigned to an argument.

Now by (21), if the doubled NP is an argument, the clitic cannot be. If the clitic is an argument, the doubled NP can only be appositional. This latter type of relation is not uncommon. For example, Jeanne 1978 argues that such appositional NPs are found in each major phrasal category in Hopi while Bresnan and Mchombo 1987 note a similar possibility for direct object NPs in Chichewa. One approach to the problem of the status of Yagua clitics with respect to argumenthood is to determine whether or not Yagua doubled NPs are in A-positions. A-positions are defined in Chomsky 1981:47 as:

'An A-position is one in which an argument such as a name or a variable may appear in D-structure; it is a potential θ-position. The position of subject may or may not be a θ-position; depending on properties of the associated VP. Complements of X' are always θ-positions, with the possible exception of idioms.'

Thus, if we can establish that doubled NPs are in A-positions (at least for VP, NP, PP), we can conclude that these NPs are arguments and therefore by (21) above, that Yagua clitics are nonarguments.

Some of the diagnostics which may be applied to determine whether a given nominal expression is an A- or A'-position are given in Table Three:

	A-positions	A'-positions
Allows bare quantifiers at D-structure	yes	no
Allows extractions	yes	no
Shows weak-crossover effects[4]	yes	no
Phonologically salient in relation to rest of clause	not usually	frequently
Pragmatically salient	not without special devices e.g. intonation, stress, etc.	usually

TABLE THREE

As indicated, we would expect results of the first three diagnostics of Table Three to be straightforward, 'yes' or 'no'. For the phonological and pragmatic tests, we can only expect that if there are differences in a given language, they will be gradient, tending to cluster as indicated. Therefore, the syntactic criteria provide the strongest and most reliable test for A versus A' positions. Consider as an example the bare quantifier test, using data from English (see Cinque 1984 for data from Italian and references to studies of this property crosslinguistically).

(22) a. John came.
b. John, he came.
c. John he-came.

For many dialects of English, (22b) is grammatical only with a pause between John and he. For other dialects, pause is not necessary, rather the pronominal he may be cliticized to the verb. This has led some (especially Givón 1976) to conclude that the cliticized pronoun in (22c) is actually a form of agreement and that the NP, John, is now the Subject (and not the Topic). However, this hypothesis cannot account for the crossdialectally valid grammaticality judgments which follow:

(23) a. { Everybody / Somebody / Nobody } came.

b. { *Everybody / *Somebody / *Nobody }, he came.

c. { *Everybody / *Somebody / *Nobody } he-came.

The bare quantifiers, everybody, somebody, nobody may appear only in A-positions. In (23a) they appear in Subject, [NP,S] position, and A-position, and the sentence is grammatical. But the sentences in (23b) and (23c) are bad. This is because in both of these cases, the pronoun is the grammatical subject, leaving only an A'-position (COMP or Topic) available for the bare quantifier at D-structure. A pragmatic account of such contrasts will not work, as the following contrasts between Standard Italian and the northern dialects, Trentino and Fiorentino show (from Rizzi 1985):

(24) *Nessuno lo conosco in questa citta. STANDARD ITALIAN
nobody him I:know in that city
'I don't know anyone in that city.'

(25) Tut l'e capitã de not. TRENTINO
everything it happens at night
'Everything happens at night.'

(26) Nessuno l'a detto nulla. FIORENTINO
nobody he has said nothing
'Nobody said anything.'

These examples show that the failure of bare quantifiers to be doubled by clitics in some languages is not a universal pragmatic restriction based, say, on definiteness contrasts between quantifiers and clitics but is a syntactic fact. The most straightforward account of the contrast in (24)-(26) above is that Standard Italian clitics are argumental, preventing co-occurring bare quantifiers from occupying A-positions at D-structure. But for Fiorentino and Trentino, the clitics are nonarguments. Rizzi (1985) shares this conclusion, arguing that the clitics in (25) and (26) are under the AGR(eement) node (following a suggestion in Belletti 1982). This allows the bare quantifiers tut 'all' and nessuno 'nobody' to occur in [NP,S] position at D-structure.

Now, applying this first diagnostic to Yagua, the evidence suggests that bare quantifiers in Yagua can be doubled, indicating that doubles are in A-positions and thus are arguments. Certain quantificational pronouns in Yagua are formed from the numeral tíí 'one' plus a suffix further specifying its meaning:

(27) a. tíí-tiy 'whoever'
one-relativizer

b. tíí 'someone, anyone'

c. tíí-quii 'one, something (animate)'
-animate

d. tíí-ki 'one' (used for jaguar)
-?

e. tii(-)taju 'all, everyone'
(-)?

Now, consider the following pair of sentences:

(28) Tíí-tiy$_j$ jiya-sara tõõ-va sa$_i$-sũũy
one-rel$_j$ go-habit jungle-dative 3SG$_i$-bite

-maa coodiy$_i$-ntiy-níí$_j$.
perfective snake$_i$-repetitive-3SGII$_j$
'Whoever goes to the jungle, the snake has bitten him/her too.'

(29) Teta vurya-a junúuy-ra
unless 1PLINC-irrealis look-inanimate

vurya$_i$-a dííy tíítaju$_i$.
1PLINC-irrealis die all
'Unless we look at it, we will all die.'

(30) Néé -tiy -múy dííy-vay-ríy
negative-conclusive-negative die-nom:PL-3PL

riy$_i$-vicha tiitaju$_i$.
3PL-live all
'If they are not dead ones, they are all alive.'

The quantifier tíítiy 'whoever' is in preverbal subject position in (28), while its morphologically related form, tiitaju 'all, everyone' is in the normal postverbal position of a doubled NP subject in (29) and (30). In spite of the gloss 'all', tiitaju is not a floating quantifier. In nondoubled contexts it appears to be restricted to subject, object, and relative head positions, i.e., A-positions, as with English everyone. On the basis of this restriction to A-positions and its grammaticality in doubling contexts, I submit that bare quantifiers in Yagua may be doubled, strongly supporting the assertion that Yagua clitics are nonarguments. Let us now consider evidence from extraction.

(31) Sa$_i$-siryi jasiy nuñú$_i$, coodiy, jááyanú.
3SGI-scurry there isúla, snake, fer-de-lance
(ant species)

tiitáyu$_i$ níí-tiy$_i$ sa$_i$-vicha-sara
all 3SG-relativizer 3SGI-be-habit

judfa súúy-ra
hurting bite-3SGINFI
'There scurried up the isula (ant species) the snake, the fer-de-lance, all who (they) are hurting, biting ones.'

Adopting D. Payne's (1985:70ff) analysis of Yagua relative clauses, the relativizer corresponding to English WH-relative words in COMP is níítiy. Now note that níítiy is doubled in (31) by sa-. This can be interpreted in two ways. Either sa- is a resumptive pronoun (a phonologically realized variable), in the sense of Sells 1984, or the relativizer binds an empty category variable (Chomsky 1981:101ff) and sa- is merely a nonargumental, nonresumptive clitic.

The resumptive pronoun hypothesis is unlikely, because it does not explain the failure of nonclitic pronouns to serve as resumptive pronouns as well, as is common in languages with resumptive pronouns. There are no structures in Yagua in which a quantifier directly binds a free pronoun:

(32) *[$_{X\ max}$...nfltiy$_{i}$...nfl$_{i}$...]

Under our present assumptions, the grammaticality of (29) and (30) versus the ungrammaticality of (32) is explained by assuming Yagua clitics to be nonarguments whereas Yagua free pronouns are arguments.

Then, since Yagua does not otherwise have resumptive pronouns, I conclude that extraction facts provide additional evidence to that already gleaned from bare quantifier constructions that doubled NP positions in Yagua are A-positions and, therefore, θ-positions. This means once again that Yagua clitics are nonarguments.

I was, not surprisingly, unable to find any examples of weak-crossover in the corpus available[6]. In any case, the evaluation of weak-crossover data requires access to native speaker intuitions about often very subtle contrasts of grammaticality. I will therefore close this section on the argument status of Yagua clitics by considering potentially relevant phonological and pragmatic data, the weaker of the diagnostics listed in Table Three.

Doubled NPs given in examples throughout this paper are not given any special phonological marking, such as pause, extra stress, intonation, etc. to set them off from the other constituents of the phrase. This is of course to be expected if doubled positions are phrasal arguments in A-positions. Phonological evidence is therefore compatible with our assertion that doubled NPs are in A-positions.

Pragmatic evidence is also consistent with this claim. If doubled NPs were in A'-positions, we might expect them to function as clarification, afterthought, or intraphrasal topics, all roles with a higher degree of pragmatic salience than subjects, objects, and other A-positions (see Dooley 1982 for a lucid discussion of pragmatic and phonological distinctions along this line in Mbyá-Guaraní). However, doubled positions are no more marked pragmatically than nondoubled positions. They are simply interpreted as the subject, direct object, etc. of their phrase. Interestingly, D. Payne (1985:201ff) asserts that clitic-doubled constructions are less salient pragmatically than nondoubled constructions with full NPs. While this is completely consistent with an analysis of doubled positions as A-positions, it would be less likely under the hypothesis that doubled positions are A'-positions. I conclude, therefore, that evidence from bare quantifiers, extraction, phonology, and pragmatics supports the analysis of Yagua doubled positions as A-positions. Consequently, I submit that Yagua clitics are nonarguments. In the following sections, I want to explore the possibility that in spite of their nonargumental status, Yagua clitics nonetheless require Case.

4 Case and word order

4.1 Introduction

To discuss Case, we must first have some understanding of what is meant by Case in GB. Case is one of the licensing conditions on well-formed expressions (Chomsky 1986a). Essentially, Case is related to θ-roles in that it is a necessary condition on the Logical Form (LF) interpretation of a θ-role: A θ-role is visible at LF if it has Case (Chomsky 1981:337). This is most obvious in 'free word order languages' such as Greek and Latin, where the morphological case-ending of a nominal expression is crucial to the interpretation of that nominal's θ-role. But, as Chomsky (1988:101ff) says, 'If all languages are essentially alike in their deeper essential nature, we would expect [languages such as, (DLE)] Spanish and English also to have a case system of this general sort.'

Indeed, like Greek and Latin, all languages impose morphological or syntactic restrictions on the relation between phrasal heads and their complements. The restriction is morphological in Greek: an NP object can only be related to a verb if it bears the appropriate case-ending, usually accusative. In English, we might say the restriction is syntactic, strict word-order relations must be maintained, e.g. subject precedes the verb, object immediately follows, etc. Notice, though, that this cannot be reduced to some sort of pragmatic 'confusion-avoidance' principle:

(33) a. John shot the dog.
b. The dog was shot.
c. *Was shot the dog.

The restriction forcing the patient of a passive into subject position, as shown by the contrast between (33b) and (33c) cannot be simply to avoid confusion as to roles, since there is no other argument around to provoke any confusion in either example. We can express the contrast rather by saying that the object in (33c) has no Case, none being available from a passive participle. It must therefore raise to subject position to receive nominative Case. Word order restrictions in English, then, are at least partly the result of Case requirements.

We can subsume Greek-type Case systems and those of the English variety under a single generalization: 'Overtly mark the relation of the complement to its head before LF.' Let us call this marking, whatever form it actually takes (whether word order, morphological endings or a combination of the two), Case. The need for this marking we shall label (following standard practice) visibility. Then, Case is intimately related to θ-role assignment (via visibility) but is nonetheless logically distinct from it.

In Everett 1986, I argue that any nominal may be visible at LF in one of two ways. It may be assigned Case (directly by the head in an A-position or via its chain, as per Chomsky 1981:333ff) or, like incorporated nouns, affixes, and many clitics, it may be visible by virtue of its appearance on a phrasal head (which itself must be licensed or visible at LF by Case-theory, X'-theory, θ-theory, etc. cf. Chomsky 1986a). These options may be formalized as in (34):

(34) Visibility:

a. Position Visibility: α is position visible if and only if for some β ($\in X^{max}$), β is Case-marked and $\beta = \alpha$ or β is a projection of α.

b. Morphological Visibility: α is morphologically visible if and only if for some γ ($\in X^{o}$), γ is position visible and γ does not exclude α.

(35) α is not excluded from γ only if it appears in a. or b.:

a. $[_{\gamma} \ldots \alpha \ldots]$
b. $[_{\gamma}[_{\gamma} \ldots] \ldots \alpha \ldots]$ (Order irrelevant)

4.2 Case assignment

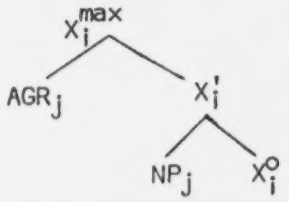
In GB, Case assignment is accomplished under government by coindexation with a Case assigner. Nominative Case is assigned by AGR, genitive Case by N^{o}, accusative Case by V^{o}, and oblique Case by P^{o}. However, the inclusion of AGR in the list of Case-assigners introduces a slight incongruency. It is the only nonhead which assigns Case. AGR is itself a constituent of INFL, the head of S (under most current assumptions, cf. Chomsky 1986b). One might hypothesize that AGR is the head of INFL and hence, ultimately the head of S, but independent motivation for such a move is not apparent. Let us say, rather, that the X^{o} head of a projection, such as INFL, may assign Case via its 'appropriate morphological constituent'. This may be the head itself or an AGR node within the head's maximal projection.

4.3 Multi-AGR languages

In Everett 1987, I argue that in Pirahã each phrasal head contains an AGR position filled by a clitic. In fact, there are many languages in the Amazon (e.g. Pacaas-Novos, Oiampi, Pirahã, and Yagua) and elsewhere which seem to manifest agreement between all phrasal heads and their complements. Let us refer to such languages as multi-AGR languages.

In such languages, it is useful to suppose that each X^{max} dominates an AGR node. This AGR node may be a morphological constituent of its appropriate X^{o} at D-structure (as in Pirahã). Alternatively, as with [AGR, INFL] and V^{o} in English and other languages, the AGR node may be discontinuous from its S-structure host at D-structure. In either situation, however, Case will be assigned internal to each X^{max} by an AGR node/morpheme. The kind of Case actually assigned will be determined by government of AGR by the head, X^{o}. The governor of AGR determines the Case AGR assigns.

(36)



The order and structure of (36) will be determined by Case assignment and θ-assignment requirements (e.g. government and adjacency), X'-theory, and language-specific peculiarities where relevant (see Travis 1984 for detailed discussion). The coindexations in (36) are forced by X'-theory. AGR assigns Case by coindexation and government to NP.

This account of general features of agreement in multi-AGR languages requires no theoretical apparatus beyond that already required for single AGR languages such as English. It is simply a logical extension of existing assumptions to a relatively unexplored domain. With this much established, let us return to consider Yagua clitics.

We concluded earlier that Yagua clitics are nonarguments, which we can represent informally as [-A(rgument)]. Also, we noted that a likely place to insert [-A] clitics is under an AGR node (Everett 1987, Rizzi 1985, Koopman 1984 and others)[8]. Let us assume for the moment, then, that Yagua nominal clitics are generated under AGR.

But if this is all there is to say about Yagua clitics, that they are nonargumental agreement morphemes, then it is difficult to see how their presence or absence would affect the syntax as drastically as we have observed. In Pirahã, for example, where clitics are nonarguments, inserted under AGR, and do not require Case (Everett 1987), clitics do not materially affect the syntax, aside from identification of <u>empty categories</u>. Their principal role is the overt registration (by coreference) of the head-complement relation.

We have already established that, like Pirahã clitics, Yagua clitics are nonarguments, [-A]. If, however, unlike Pirahã clitics, Yagua clitics required Case, then it would be natural to expect that they would produce a more obvious effect on the syntax than non-Case-requiring clitics, since Case is so tightly restricted by configurational and linear requirements.

Furthermore, under this assumption, any clitic-doubling construction would contain a Case-conflict: the clitic and its double would compete for a single Case assigned by the head[9]. This Case conflict would rule all clitic-doubling structures ungrammatical unless there were some means in Yagua grammar for resolving this conflict[10]. I am going to argue in what follows that in fact a solution to this conflict is available from independent, universal principles. Let us begin by considering an analysis of clitic-doubling in VP[11].

4.4 Case assignment and doubling in VP

4.4.1 Direct objects. Recall that doubling of a direct object is optional when the direct object is right-adjacent to the verb, required elsewhere.

(37) a. Sa-púúchu(nîí$_j$) Anita$_j$.
3SGI-carry(-3SGII) Anita
'He carries Anita.'

b. Sa-púúchu Pauro *(-nîí$_j$) Anita$_j$.
'Paul carries Anita.'

(38) a. Anita ra rumiy(ra$_i$) buyaa$_i$ yi-iva.
Anita irrealis spill(-3SGINII) manioc:beer 2SG-DAT
'Anita will spill manioc beer on you.'

b. Anita ra rumiy yi-iva *(-ra$_i$) buyaa$_i$.

The asterisk preceding the parentheses indicates that the material inside is obligatory.

There are two problems we must deal with. First, how is Case assigned? Second, why must the clitic appear when the object is nonadjacent to the verb? Consider the first question.

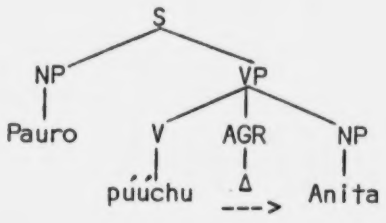
The requirement that objects follow the verb could be expressed in terms of Case:

(39) Assign Case to the right.

This is natural given that Cases are generally assigned in a single direction. Another near universal restriction on Case assignment is that

the Case assigner and assignee be adjacent (Stowell 1981). Given these notions, consider a possible structure for VP in Yagua:

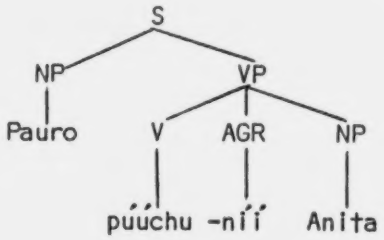
(40)



(→ = accusative Case assignment)

In simple SVO structures, then, accusative Case is assigned to a right-adjacent NP. If, as we are supposing, Yagua is a multi-AGR language and AGR optionally dominates a clitic, then an alternative, more detailed expansion of (40) would be:

(41)



In this configuration, but in no other, the Set II clitic is optional. Let us annotate this as in (42):

(42) When a Case is assigned to a right-adjacent NP, the clitic is optional; otherwise, it is obligatory.

What might account for (42)? As Safir (1981) has argued, in most languages, it is required that Cases be <u>realized</u>[12]. The realization of Case is its phonetic implementation, that is, how it comes to be heard. As a set of morphosyntactic restrictions between heads and complements, Case is realizable, as we have noted, via linearity, morphological marking, and/or other ways (such as, perhaps suprasegmental devices). Suppose that in Yagua, Case realization must obey (43):

(43) <u>Yagua Case Realization</u>: A complement must appear right-adjacent to its Case assigner at S-structure.

Suppose, too, that the clitic, even though we are assuming that it also needs Case, may assign Case to the object. This would account for all of the object-doubling facts we have seen to this point.

When the object is right-adjacent to the verb, the verb may assign Case to it directly, by (43). Alternatively, the clitic could assign Case to the object (we have not yet seen how, but we will). Then, whether the clitic does or does not appear, (43) is satisfied. But now, when the object is not adjacent to the verb, it can only satisfy (43) if doubled by a clitic to its immediate left which assigns it Case. If this could be made to work in a satisfactory manner, then the principles of Case assignment (42) and Case realization (43) will account simply for placement and obligatoriness of direct object clitics.

The problem of course is how the clitic could assign Case to the object. Recall our assumption that Yagua clitics need Case. An object clitic will have only one potential Case-assigner, the verb. But, if the verb assigns accusative Case to the clitic, the doubled object will fail to receive Case. Thus it seems difficult to reconcile our assumption that clitics need Case with the facts[13].

And yet, there is one possible analysis. As I noted in (34), there are two ways that a nominal can be visible: by receiving Case or by appearing on a licensed X^0 category. To say that a given expression requires Case is to say that (i) it must be assigned Case and that (ii) it must be visible at LF. Curiously, perhaps, both of these requirements may be met without necessitating that the expression actually have Case at S-structure. An expression could be assigned Case, acquire morphological visibility, and subsequently allow its Case to be reassigned. Of course, the only kinds of expressions which could acquire morphological visibility are those which may attach to X^0. Since Yagua lacks a rule of noun-incorporation (which I take to be an arbitrary characteristic; some languages have this rule, others do not), this leaves only clitics and affixes, which I will lump together for now as dependent terms.

Yagua objects may be separated from the verb either by a doubled subject or by movement/placement of the object to the rightmost boundary of the VP. Then, Case assignment, Case realization, and visibility can account for all the facts.

When the object is right adjacent to the verb, the clitic need not occur. If it does, then it must attach to the verb (which is independently visible via predication, as per Rothstein 1983; Fabb 1984). This frees up its Case for reassignment (subject to the same restrictions as initial assignment). When the object is not adjacent to the verb, it cannot receive Case, assuming that Case is assigned after movement, nor have its Case realized via right adjacency to the verb. But if the verb's objective Case were assigned to the clitic, this problem would be eliminated. Assume the general rule Affect- (Lasnik

and Saito 1984), which allows us to move or otherwise change the D-structure characteristics of a nominal expression in any way we please, subject to independent restrictions. Then Case principles will attach the clitic to a X^0 category left-adjacent to its double and reassign its Case to its double (after acquiring morphological visibility) whenever the double is not right-adjacent to the verb[14]. This gets us every fact we have noted to this point:

1. Obligatoriness of the clitic when no double is present (Case must be assigned and realized;
2. Optionality of clitic when double is otherwise right-adjacent to the verb (Case realization is satisfied with or without the clitic);
3. Obligatoriness of clitic when double is not right-adjacent to the verb (Case realization can only be satisfied when the clitic is present);
4. Clitic cannot attach to its double (the double is not independently visible);
5. Clitic must attach to the immediate left of the double (Case is assigned to the right);
6. Set II clitic attaches to right edge of host, rather than left edge (direction of Case assignment and adjacency requirements)[15].

We get these facts without having to assume that clitics are inherently different from affixes (for example, tense/aspect markers do not move because they are not involved in Case assignment, not because they belong to a different type of morpheme class. See section seven.) nor that there is some sort of 'wrong-way' cliticization going on (as suggested in Payne 1983). The facts all follow from independent, syntactic Case principles. As we now see, this analysis applies straightforwardly to doubling of indirect objects.

4.4.2 Indirect objects. D. Payne 1985:30 and Payne and Payne 1988 note several facts about indirect object doubling in Yagua. However, what is truly interesting is that all these facts can be expressed by the generalization that the object closest to the verb need not, but may be, doubled (as we have already noted in the context of direct objects), while the object farthest from the verb must be doubled. When the verb is bitransitive, each subcategorized object must be overtly referenced, either through a clitic or an argument NP.

(44) Sa-saay*(-nii)*(-ra)
Sub-give-Iobj-Dobj
'He gives it to him.'

(45) Rodrigo saay-(nii$_i$) ravichu$_i$-ray.
Rodrigo give-3SGII rock = 1SG
'Rodrigo gives me the rock.'

The clitics are required in (44) for the same reason noted for direct objects; the Case must be realized.

(46) Sa-dáátya-ny-(nîi$_i$) Antônio$_i$ *(-ra$_j$)
3SGI-know-transitive-3SGII Antônio-3SGINII

niquee-jada$_j$.
talk-infinitive
'He teaches Antônio the word.'

(47) Yi-a-saay-ray$_i$ tîîquii quivää$_j$
2SG-irrealis-give-1SGII one fish
'Give me one/a fish.'

Example (47) is interesting because it shows that a clitic, -ray '1SGII', does not count in determining adjacency. A nonclitic cannot intervene between the verb and a nondoubled direct object, as we have seen in numerous examples above. That clitics do not count for adjacency is expected if they Chomsky-adjoin to their host, as in (48):

(48) a. $[_V \cdots]$ $[_{Clitic} \cdots]$

b. $[_V\ [_V \cdots [_{Clitic} \cdots]]$ [16]

4.5 Definiteness

Previous studies of Yagua have pointed out that NPs doubled by Set II clitics are interpreted as definite. This is illustrated in examples such as (17) and (18) above, repeated here:

(49) a. Jir -rimiy-maa ray-rááva-ta
2SGI-spill-perfective 1SGI-poison-partitive
'You spilled part of my poison.'

b. *Jir-rimiy-maa-ra ray-rááva-ta

3SGINII

As is shown in (49), a noun in partitive Case cannot be doubled. This restriction does not apply to other Cases. Now, compare this with other examples like (50):

(50) a. Sa -jatu buyaa.
3SGI-drink manioc:beer
'He drinks manioc beer.'

(either generic or definite interpretation allowed)

b. Sa-jatu-ra$_i$ buyaa$_i$.
'He drinks the manioc beer.'

In (50), we see that a doubled NP must be definite (although a nondoubled NP is not obligatorily nondefinite). This explains the

restriction against doubling partitive NPs, as in (49b), since partitive Case is inherently nondefinite and thus incompatible with doubling (because this produces definiteness). The fact that a doubled NP must be definite is not uncommon crosslinguistically (Givón 1984:371ff).

I want to suggest an account of this phenomenon which, although highly speculative, nonetheless indicates that the Case-based analysis of Yagua clitic doubling may be on the right track. This proposal is based on work of Adriana Belletti (1988) in which it is argued that the inherent Case of certain verbs is partitive.

In GB, an inherent Case is one which is associated with a restricted range of θ-roles. Such Cases are only assigned when the Case-assigner also assigns a θ-role to the Case-assignee. Genitive and oblique Cases are inherent Cases, for example. (This restriction is known as the Uniformity Condition; see note 9 above. Belletti 1988 and Chomsky 1986a discuss inherent Case in detail.)

Yagua is like many Amazonian languages in that it lacks definite/indefinite articles, or any other obvious formal device for distinguishing definiteness. Thus, it would not be difficult to imagine that if a device existed independently that could be exploited to mark definiteness, the language might take advantage of this.

Now, in GB there are only two Cases which are not inherent Cases, nominative and accusative, which GB labels structural Cases. That is, only these Cases may be assigned even when the assigner does not θ-mark the assignee. Recall that in Yagua, Set I clitics double nonobject NPs. Thus, they may appear on Ns, Ps, and AGR. Although AGR assigns a structural Case, nominative, N and P assign only inherent Cases, genitive and oblique. Thus, Set I clitics are not unambiguously associated with structural Case, but also may reassign/receive inherent Cases. Set II clitics, however, are unambiguously associated with structural Case. That is, since Set II clitics only double objects of V, then any time a Set II clitic appears, a structural Case may be assigned. Because a verb also θ-marks its object, it may assign either a structural Case or an inherent Case, an option unavailable for any other phrasal head.

Now let us suppose that Yagua has exploited these facts in the following way. A verb may assign either an inherent Case, which by hypothesis will be the nondefinite partitive Case, or it may assign a structural Case (accusative). Let us assume that in Yagua the distinction between these Cases will be realized only in the definiteness of the object (verbal inherent Case = nondefinite; verbal structural Case = definite). When there is no Set II clitic, the verb may assign either Case, so nondoubled objects may be either definite or nondefinite. However, since Yagua clitics are nonarguments and since inherent Case may only be assigned to arguments, only structural Case may be assigned to a Set II clitic. But this means that only structural

Case will be available to be (re)assigned to the double of a Set II clitic. To express definiteness via Case, Yagua need only mark inherent V Case as nondefinite (according to Belletti 1988 this is a property of Universal Grammar (UG)) and structural Case as definite (a logical extension of a contrast potentially available in UG). Then any NP doubled by a Set II clitic will receive structural Case and, therefore, a definite interpretation.

This will not apply to Set I clitics because they are not unambiguously associated with a structural Case assigner. The Yagua child would have no clear data available to learn this distinction, since Set I clitics with Ns and Ps would have to receive an inherent Case, while Set I clitics with AGR would receive structural Case.

The restriction against assignment of inherent Cases to nonarguments may be enforced for Set II clitics at all levels of the grammar since verbs may assign either inherent Case or structural Case. But for Set I clitics, since N^{o}s and P^{o}s only assign inherent Case, this could only be enforced subsequent to the clitics' acquisition of morphological visibility. I assume therefore, that for Set I clitics at least, the Uniformity Condition is only checked at Logical Form.

While this account is incomplete, it is sufficient to show that the present analysis is not only compatible with these facts, but is able to point to an explanation of this set of facts in terms of Case in UG. Since this same Case subcomponent of GB has informed our entire analysis, this looks like a very promising avenue of inquiry. Therefore, I conclude that definiteness facts offer intriguing potential support for my analysis of Yagua clitics.

4.6 Summary

We can sum up the facts about object doubling:

(51) a. Case must be realized in Yagua.
b. Case is always assigned to a right-adjacent NP.
c. Case realization is always the appearance of the Case assigner left-adjacent to the Case bearer at S-structure.
d. Yagua Set II clitics are [-A, +C].
e. As nonarguments unambiguously associated with V, a structural Case assigner, Set II clitics are assigned structural Case.
f. Therefore, NPs doubled by Set II clitics must receive a definite interpretation.

4.7 Unaccusative subjects

This analysis also applies across the board to that class of intransitive subjects which are doubled by Set II clitics, as in (52)-(53) (see note 17 also):

(52) Machîturu-numaa(nîî$_1$) Antônio$_1$.
teacher-now(3SGII) Antônio
'Antônio is now a teacher.'

(53) Machîturu-numaa-nîî.
'S/he is now a teacher.'

T. Payne 1985:105ff specifically relates the class of intransitive verbs taking Set II clitics, which he labels S_o clauses, to unaccusative verbs, referring to Relational Grammar as his source. He notes that for Yagua, as for many languages, there exist 'stative/nonvolitional' subjects which are treated 'morphosyntactically just like transitive objects'. Payne further remarks that such subjects are often associated with verbs of locomotion and being (as in equative or predicate nominal constructions), two classes commonly associated crosslinguistically with unaccusatives (Burzio 1986; Davies 1986). He also observes that many verbs only optionally take S_o subjects and may instead take S_a subjects (i.e., those with Set I clitics, cf. section 4.8.1 below), depending on their discourse function. This will not concern us here, however.

Given the lack of detailed information available on these verbs, and their close correspondence semantically and syntactically to unaccusative verbs, I will assume here that S_o clauses are just unaccusative clauses. When the subject is postverbal, it is in the VP object position and thus may (optionally, for the same reasons of Case realization noted above) be doubled by a Set II clitic. When unaccusatives appear postverbally, the preverbal position is obligatorily filled by a 'presentational' word (Payne's terminology) such as mûûy 'there'[17].

4.8 Case, subjects, possessors, and obliques

4.8.1 Subjects. Let us begin this section by reviewing the facts to be accounted for. We want to know (i) why doubling by a Set I clitic forces the double to appear to the right of the head, whereas nondoubled NPs appear to the left; (ii) why Set I clitics undergo word-internal processes such as Vowel Harmony, while Set II clitics do not; (iii) why the host of a Set I clitic, like that of a Set II clitic must precede the double; (iv) why Set I clitics only attach to the phrasal head. I argue in this section that the analysis developed for Set II clitics will account for all of these facts with the additional information that Set I clitics appear to the left of their host, a fact for which there is a reasonable historical explanation (cf. section seven below).

Consider once again examples of subject doubling:

(54) a. Pauro púúchu Anita.
'Paul carries Anita.'

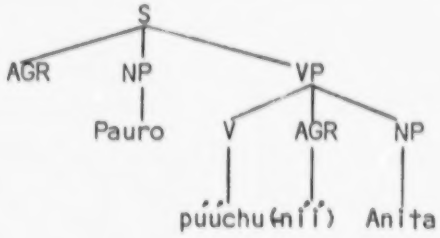
b. Sa-púúchu Pauro-ní́ Anita.
'Paul carries Anita.'

(55) a. Tomása$_i$ dííy yi$_i$-íva.
Tom see COR-DAT
'Tom sees himself.' (lit: 'to himself')

b. Sa$_i$-dííy Tomása$_i$ yi$_i$-íva.
'Tom sees himself.'

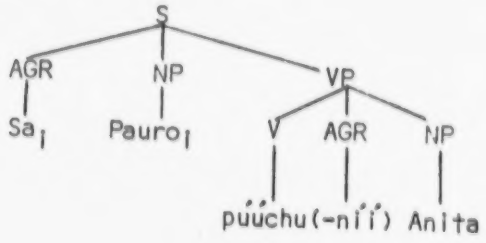
If we treat Set I clitics as AGR and assume that Case in Yagua is assigned to the right, just as we did for Set II clitics, then we derive a structure along the lines of (56):

(56)



AGR then assigns Case to the right both in VP (accusative Case) and under S (nominative Case). Now, if a clitic is inserted under sentential AGR, as in (57), since the clitic needs Case, a Case-conflict arises:

(57)

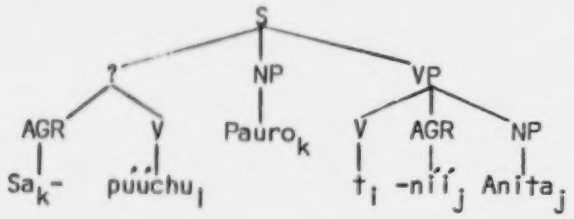


Both the clitic sa and the NP Pauro are competing for a single Case (nominative). If AGR assigns Case to sa, then Pauro fails to receive Case and vice-versa. As with Set II clitics, to resolve this conflict, we need recourse to morphological visibility, so that the Set I clitic may reassign its Case to the subject.

Recall from definition (34) above that an expression can only acquire morphological visibility if it is attached to an independently visible host. Thus, the clitic cannot attach to its double since this would fail to resolve the Case conflict. Fortunately, a natural host for AGR is available, the verb. It is a well-known fact that AGR generally must occur on the verb by S-structure. Also, since the verb is visible independently via predication, the Case-conflict in (57) can be resolved by attaching the clitic to the verb, either by lowering the clitic or by raising the verb.

The first option, clitic-lowering, must be rejected since then the clitic's nominal Case would still not be available for assignment to the [NP,S] position. This is because Case in Yagua must always be assigned rightward, as we have seen. But if the clitic lowered, its double would occur to its left and Case could not be assigned. The only option, therefore, is for the verb to raise to INFL/AGR position. As argued in Chomsky 1986a and Travis 1984, V-to-INFL movement is a rather common operation crosslinguistically. This would derive the correct orders:

(58)

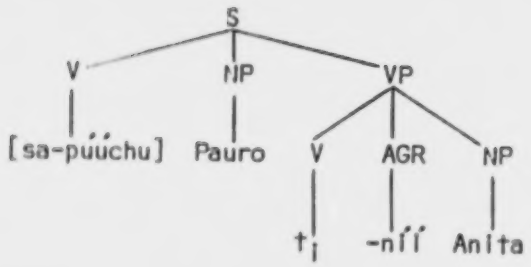


Sa may now attach to the verb, freeing up its Case for reassignment to the immediate right. But there is still a problem. The verb intervenes between the clitic and the [NP,S] position. Thus, the Case assigner and Case assignee are not adjacent. But if the clitic were actually included in its host, then it would in fact be adjacent to its double. This is so since it would then in effect be part of the verb. We defined exclusion in (35) above. Inclusion may be defined as:

(59) Inclusion: α includes β only if every segment of α dominates β . (Chomsky 1986b:7ff)

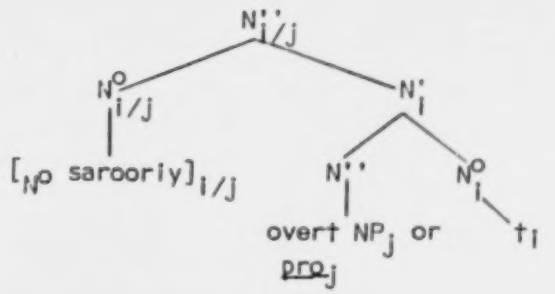
This will mean that when V moves to AGR it does not Chomsky-adjoin to AGR but, rather, sister-adjoins to AGR, which will subsequently cliticize to V (on the attachment of inflectional morphemes in general, the reader is referred to Emonds 1985:243ff). Thus the structure of (58) will be in fact as in (60), subsequent to cliticization:

(60)



A final comment will complete our analysis of subject-doubling. Since the clitic bears an index prior to cliticization, this index will percolate up to the maximal projection of its containing X^o. This is a consequence of the fact that the index of the clitic just is an index of X^o, since the clitic is part of X^o. For example, the structure of a NP with doubling would be as follows (NP structure is discussed in the next section):

(61)



As we see in section 9 below, this makes some interesting predictions which strongly support the present analysis[18].

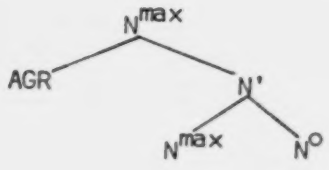
We thus account for why (i) Set I clitics are more closely attached to their hosts than Set II clitics (otherwise they would be nonadjacent to their double for questions of Case assignment), undergoing word internal processes with their host; (ii) SVO is obligatory without a clitic but (iii) VSO is obligatory with a clitic. Let us now turn to consider doubling of possessors and adpositional objects. The analysis just developed for subject doubling applies virtually without modification within NPs and PPs.

4.8.2 Possessor doubling. If Yagua indeed is a multi-AGR language, then it will have the following properties: (i) each X^{max} projection will dominate an AGR node; (ii) agreement morphemes will be inserted under AGR; (iii) Case may be assigned by AGR, as the 'appropriate morphological constituent' of the head. As with AGR, INFL, and V in many languages, (iii) may apply whether or not AGR and the head of its phrase form a continuous constituent at D-structure.

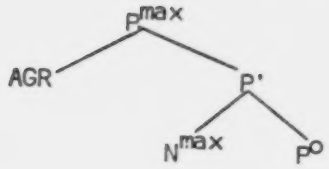
We assumed that nominative and accusative Cases are assigned to the right by AGR. I maintain this assumption here. This will mean that AGR appears to the left of the NP complement in NPs and PPs. If we also assume that trees are always binary, then structures like (62a) and (62b) are derived:

(62)

a.



b.

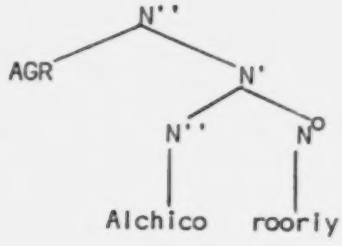


If we do not assume that all trees are binary, the configurational relations will not be so straightforward. For the sake of this discussion, I will assume binary trees.

Consider a NP with a nondoubled possessor:

(63) a. Alchico rooriy.

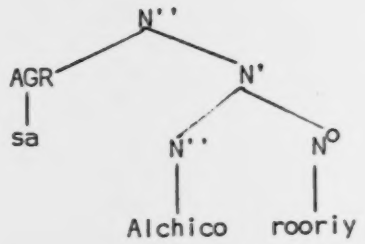
b.



<u>Alchico</u> receives genitive Case from AGR. Genitive Case is thus assigned to the immediate right, maintaining uniformity of assignment

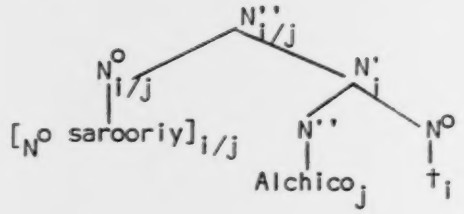
for nominative, accusative, genitive Case. But now suppose a clitic is inserted:

(64)



Sa must acquire morphological visibility or the Case conflict between sa and Alchico will produce ungrammaticality. This can be resolved in the same way as nominative Case. Allow N^{o} to move to AGR and formally include AGR. Then sa's genitive Case will be freed up and able to be reassigned to Alchico via the standard means of government, coindexation, to the immediate right:

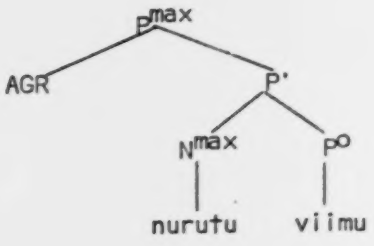
(65)



The same analysis may be used for doubling in PPs.

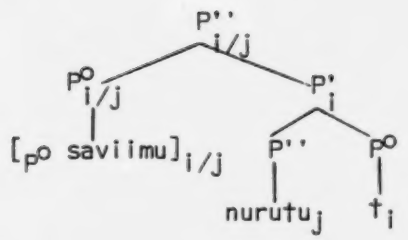
4.8.3 Postpositional phrases. I assume the following derivations for doubling in PPs. The reader is referred to the preceding section for details.

(66)



(67) a. Sa-viimu nurutu.

b.



4.8.4 Conclusion. We have been successful in accounting for a wide range of intriguing facts of Yagua clitic doubling. Further evidence in favor of the hypothesis defended above is adduced in the next section from reflexives and coreference forms. Because this entire analysis depends on the assumption that Yagua clitics are [+C,-A], its success is to be understood as supporting this type of classification. The crosslinguistic classification of clitics according to the parameters of Case and Argumenthood derives from findings in Everett 1986 and 1987. The interaction of these two features predicts exactly four types of clitics crosslinguistically. The reader is referred to the references just cited for details. The upshot of this is that the study of Yagua clitics gives important new information to studies of clitic types allowed in UG.

5 Reflexives

A crucial component of the above analysis is that the index of Set I clitics must percolate to the X^{max} projection of their host. This is motivated by Case assignment and evidenced by the degree of attachment of Set I clitics[19]. However, if the index of the clitic does indeed percolate to the X^{max} node of its host, then since the index of the clitic = the index of the double, NPs doubled by Set I clitics should have an extended c-command domain, namely, that of their containing phrase (given the indexings in the preceding examples). Possessors and obliques will therefore c-command the same material as their dominating (maximal) PP or NP[20]. Since index percolation is not motivated for Set II clitics, object NPs should not have the same c-command domain as their VP. If all indices are referential indices (Chomsky 1981; Williams 1980), these predictions should be testable by the behavior of the relevant elements relative to the binding theory.

For example, can possessors and oblique NPs serve as antecedents to elements outside of (but within the c-command domain of) their dominating NP or PP? The expectation is that they can only if they are doubled[21].

A further bit of introductory comment which must be covered is what is actually bound by the antecedent. Recall that Yagua clitics are not arguments. This means that the reflexive clitic is not itself the argument bound by an antecedent (or θ-assigned by its phrasal head). I will assume for purposes of discussion that clitics always license empty category (ec) arguments in the position of their double when no overt NP is present (this is also required by the θ-criterion since the clitics cannot themselves receive the θ-role assigned obligatorily by the verb). For nonreflexive clitics, this ec will be pro ([+pronominal, -anaphor] using the features of Chomsky 1982). But for reflexive clitics, this ec will be [-pro, +an, +argument], that is the nonovert counterpart of lexical reflexives such as himself, themselves, etc. The possible need for such a category is mentioned in Everett 1987 and in Saxon 1986. I am assuming in fact a structure like (68) for clitic doubling constructions:

(68) ...clitic$_i$...NP$_i$/pro$_i$...
(Where clitic and NP/pro refer to a single θ-role.)

The NP/pro in (68) is the double of the clitic, as defined above.

There are two types of anaphoric (or coreferential to use T. Payne's (1985:46ff) terminology) clitics in Yagua, corresponding to Set I clitics and Set II clitics, respectively.

According to T. Payne 1985:44ff, the Set I coreference clitic, jiy-/yi-, indicates (obligatory) coreference between the argument so marked and another argument which can be doubled by a Set I clitic

(possessor, oblique, and subject). The Set II coreference clitic, -yù, indicates coreference of an object (again, though, the only possible antecedents are possessors, obliques, and subjects). jíy-/yí- is most commonly used to refer to possessors. Examples of jíy-/yí are:

(69) [$_S$Núúdya$_i$-jiya-numáá [$_{NP}$jíy$_i$-rooriy-mu-jù]].
1PLEX-go-now COR-house-LOC-DIR
'We are going to our house.'

(70) [$_S$Sa$_i$-púúchi-níí Anita [$_{NP}$jíy$_i$-rooriy-mu]].
3SG-carry-3SGII Anita COR-house-LOC
'He carries Anita into his house.'

These first two examples show the binding of jíy-/yí-, itself doubling a NP object, by the subject. The next two examples show binding of this morpheme from arguments of NP and PP, respectively.

(71) [$_S$[$_{NP}$Sa$_i$-rooriy pro$_i$-mu] jíy$_i$-púúchi-níí Anita].
3SGI-house LOC pro COR-carry-3SGII Anita
'In his$_i$ house he$_i$ carries Anita.'

(72) [$_S$[$_{PP}$Naada$_i$-jinchaju pro$_i$]
3DL-upon pro COR-lie

jíy$_i$-tiryo sa-viimu koodiy].
3SGI-inside snake
'(Those) two lie upon each other in the snake.'

In T. Payne 1985:44ff, the principal syntactic restriction on coreference is that the antecedent must precede the double. This is not sufficient, however, since VP objects may not be antecedents even when they precede. Given the device of index percolation and the fact that it does not occur with direct objects, it seems that an additional condition is necessary: the antecedent must c-command the anaphor[22].

(73) Conditions on antecedents of anaphors in Yagua:
(i) The antecedent must precede the anaphor;
(ii) The antecedent must c-command the anaphor.

However, it may be possible to eliminate condition (73i), precedence. In all the examples where jíy-/yí- is bound by a preceding possessor or oblique, this argument has arguably been fronted for pragmatic effect, such as focus. Moreover, if we assume that focus is accomplished in Yagua by adjunction to S, then we will have the following inference:

1. A subject c-commands everything in S because it is immediately dominated by S.
2. A preceding possessor has been adjoined to S (note that it is

certainly out of its D-structure VP position). Therefore, its c-command domain is also S.
3. Then, since there is no evidence that a preceding NP may be an antecedent if it fails to c-command the coreference clitic (given index-percolation), we may restate (73) as (74):

(74) Condition on Yagua coreference: The antecedent must c-command the anaphor.

But (74) is obviously just a fact about Universal Grammar. This means that we have eliminated the need for any language specific statement about Yagua coreference. It is all explainable via Case theory. This would also explain why the coreference clitic may not precede its antecedent: the antecedent would no longer c-command the anaphor. That these otherwise surprising facts are actually predicted by the present analysis is to be understood as strongly supporting this analysis. A potential problem for this analysis is raised, however, when we consider examples such as (75) and (76):

(75) Sa_i-jutay-jásiy $yí_i$-a jiyá.
3SGI-say-PROXI COR-IRR-go
'He_i said (that) he_i will go.'

The coreference clitic in (75) doubles the subject position of an embedded, indirect speech clause. This would appear to violate the Binding Conditions, since it seems to allow an anaphor to be unbound in its governing or binding category, S. However, it is crucial to note that the coreference clitic jíy-/yí- may only double subjects of tensed clauses in indirect speech. It has been argued by some (cf. Kayne 1983; Picallo 1984, and others) that certain subcategorized embedded clauses may be coindexed with the matrix verb. If we assume that indirect speech clauses are subcategorized by and coindexed with the matrix verb jutay 'say', then the binding category for indirect speech clause subjects will be the matrix clause. This follows under the notion of binding category defended in Chomsky 1986:171ff:

(76) Binding Category:
Where I is an indexing, β a domain, α an anaphor, and γ a governor of α.
For some β, α must be bound in β, where β is the least CFC containing γ for which there exists an indexing J Binding Theory (BT)-compatible with (α, β).
I is BT-compatible with (α, β) if α is bound in β under I.
β is a Complete Functional Complex (CFC) if all the grammatical functions compatible with a head dominated by β are coindexed with β. (See also Johnson 1987:354).

Let us assume that the CFC for the coreference clitic on the subject of the indirect speech clause will be the main clause, since indirect speech clauses are coindexed with the matrix verb. This will

bring indirect speech clauses into line with the rest of the analysis, needing no special stipulations[23].

Then, all facts regarding the binding of the Set I coreference clitic are predicted by the interaction of Case theory and Binding Theory. Let us turn now to consider the reflexive clitic, -yù.

In most dialects, there is no morphological distinction between nonanaphoric and anaphoric clitics for first and second persons. The standard form of the Set II clitic may be used for reflexive or nonreflexive readings. In these dialects, the reflexive clitic, -yù is used only for third person singular. In other dialects, -yù is used for third person singular and first and second persons dual and plural. For example[24]:

(77) Sa$_i$-jinúúy pro$_i$-yù$_i$ pro$_i$.
3SG-look:at-3SG
'S/he looks at her/himself.' (all dialects)

(78) Jiryey$_i$-jinúúy pro$_i$-yù$_i$ pro$_i$.
2PL-look:at-2PL pro-REFL
'You look at yourselves/each other.' (upriver dialects)

For those dialects which use standard forms of Set II first and second person clitics as anaphoric, I will assume for purposes of this discussion that anaphoric and nonanaphoric first and second person clitics are lexically distinct homonyms. I will focus the discussion on -yù since the same analysis will apply to all anaphoric clitics.

Any Set I clitic double may be an antecedent for -yù. (Again, the prediction is that if the NP is not doubled, then only subjects may serve as antecedents. No examples are found in the corpus I have been able to consult that would cast any doubt on this. Some important tests, then, await future work with native speakers.) Any c-commanding NP/ec may antecede -yù, even an ec doubled by the Set I coreference clitic jíy-/yí- (as in (82)):

(79) Sa$_i$-junumívay-síy-yù$_i$.
3SGI-paint-PASTI-REFL
'He painted himself.'

(80) Sa$_i$-jumutyo jííta naana-daa-nú-yú$_i$.
3SGI-answer discourse particle 3DL-little-person-REFL
(i) 'Her son answered herself.'
(ii) 'Her son answered himself.'

Example (80), as noted by T. Payne 1985:46, is ambiguous. The reflexive may be interpreted as taking either the possessor ec or the NP head as its antecedent. The discourse particle jííta (apparently

signalling a 'thematic break', T. Payne 1985:6ff) does not affect the reflexive reading at all.

(81) Sa_i-rooriy-mu sa_i-juva-sĭy-yù$_{i/j}$.
3SGI-house-LOC 3SGI-strike-PAST1-REFL
(i) 'In his$_i$ house, he$_j$ struck him$_i$.'
(ii) 'In his$_i$ house, he$_j$ struck himself$_j$.'

(82) Sa_i-rooriy-mu $jĭy_i$-súúy-yù$_i$.
3SGI-house-LOC COR-bite-REFL
'In his$_i$ house, he$_i$ bit himself$_i$.'

These facts about possible antecedents for reflexive provide striking confirmation of the notion of index-percolation which results from the Case-based analysis of Yagua clitic-doubling proposed in section four above. Before closing, it would be useful to show clearly that objects cannot be antecedents. Both D. Payne 1985:152 and T. Payne 1985:46ff state this prohibition explicitly, providing examples like (83):

(83) Sa_i-suuta-nĭĭ$_j$ Anita$_j$ $jĭy_{i/*j}$-rooriy viimu.
3SGI-wash-3SGII Anita COR-house inside
'S/he washes Anita inside his/her house.' (house = Anita's)

Such examples are out because the object NP Anita fails to c-command the anaphor jĭy-.

I would like to turn now to consider some theoretical and typological implications of this analysis.

6 Underlying versus basic word order

6.1 Underlying word order

This analysis has assumed that the underlying word order of Yagua is SVO. It is worth considering the alternative hypothesis that VSO, the other frequently occurring word order, is underlying rather than VSO.

The hypothesis that VSO is the underlying word order fails to derive the ungrammaticality of (84) below in any enlightening way:

(84) *S clitic-V O

The VSO hypothesis fails here because (i) it can only stipulate that clitics and their doubled NPs may not co-occur preverbally; (ii) even if VSO is underlying it still has to allow for an alternate word order or derivation in which the subject appears preverbally, due to the frequent occurrence of SVO order.

In the SVO analysis, the clitic-VSO alternate order is a straightforward consequence of Case theory, as we have seen. But the VSO hypothesis has no way of deriving these facts. If it posited, for example, a rule of subject fronting, it would have to resort to stipulation to avoid the order in (84). A stipulation such as 'clitics must precede their doubles' might cover the facts but it is hardly more than a restatement of them and therefore much less desirable than an explanatory account.

Another stipulation might be that there is only one preverbal position available, to be filled either by the clitic or by a NP subject. But this depends on the highly dubious assumption that Set I clitics occupy NP positions. We have already seen that Set I clitics are part of the verb since, among other things, they undergo word internal Vowel Harmony with the verb. Thus, clitics do not occupy full word positions (A or A' positions in GB). But if this is so, the VSO analysis is hard-pressed to explain why they and full NP subjects cannot both occupy preverbal position. Moreover, as examples like (72) above show, clitics and other phrases, e.g. PPs may in fact occupy preverbal position simultaneously. This renders the restriction against simultaneous co-occurrence of clitics and subject NPs in preverbal position a stipulation at best.

Another difficulty for the VSO analysis is that, without additional assumptions, it predicts the absence of subject-object asymmetries since it implies that there is no VP underlyingly (see Chung 1983 for discussion of the implications of claiming that VSO is an underlying order). Yet we have seen clear evidence that Yagua does have the type of subject-object asymmetries we would expect if it had a VP; (i) the failure of direct objects to serve as antecedents; (ii) the entire Set I versus Set II clitic class division, which can be credited simply to a VP: Set II clitics double NPs dominated immediately by a projection of V (transitive objects and unaccusative subjects), while Set I clitics double all other NPs; (iii) unaccusative subjects differ from other subjects in being doubled by Set II clitics, exactly what is expected by the theory of unaccusatives (Burzio 1986) if Yagua has a VP; (iv) theory internal considerations also argue for the existence of a VP node: θ-role assignment arguably depends on a VP-internal versus VP-external distinction universally (cf. Williams 1980, Emonds 1985, Safir 1987).

Thus, what the SVO-as-underlying analysis gets without further ado, a VSO analysis can only stipulate or ignore (the arguments for a VP). I conclude, therefore, that the evidence strongly supports the assertion that SVO order is underlying and that this assertion is conceptually and empirically superior to the VSO hypothesis[25].

Before concluding, however, we need to consider the single syntactic argument given (Payne 1986) that VSO is basic. Payne 1986 observes that fronting the object NP into clause initial position only produces structures of the type O clitic-VS, never *OSV. Payne claims

that if SVO were 'basic', then an otherwise unmotivated rule of Subject Movement to postverbal position would be needed, triggered by Object Fronting, to rule out *OSV. Various potential responses come to mind here. For example, recall that Case Realization, (43) above, requires all arguments to be right adjacent to their Case assigner at S-structure. Let us break this into its two components: (i) all arguments must be adjacent to their Case assigner at S-structure; (ii) all arguments must be right adjacent to their Case assigner at S-structure. Assume that (i) holds for all arguments, but that (ii) holds only for sentence internal arguments. A rule such as Object Fronting (highlighting or topicalizing the object) may violate the finer constraint of directional adjacency, (ii), for pragmatic effect but not the stronger, more general nondirectional adjacency requirement, (i). Then, when the object raises to S initial position, the verb (its Case assigner) must also raise (Set II clitics may not leave the VP and hence could not be used to save the construction). If V Raising, a sentence internal rule, requires sentence internal motivation, then a Set I clitic must appear under AGR, forcing V Raising, simultaneously satisfying the adjacency constraint on the object (which holds, presumably at S'). This will allow only the order O clitic-VS, just as desired.

An alternative account would be that AGR must be governed when it does not dominate a clitic (due to the Empty Category Principle). By hypothesis, an object (or any other material, cf. examples like (72) above) moved to COMP, the governor of AGR (cf. Chomsky 1986a), would produce a branching structure, blocking government of AGR by COMP. Then, when COMP is branching, AGR may not dominate an ec and thus must be filled by a clitic, forcing V Raising and deriving the desired orders (I assume here that AGR is never null but always dominates a clitic or an ec). It would take more space than can be justified here to argue conclusively for either of these hypotheses. The point to be made, though, is that the argument raised by Payne 1986 for VSO as underlying is not a strong one. At the same time, the arguments for SVO are.

6.2 Basic word order

In spite of the above arguments on underlying order, D. Payne 1985 offers rather convincing arguments that the pragmatically unmarked word order in Yagua is clitic-VSO. For example, in Yagua discourse, SVO order is always used to introduce new participants, thematic changes, or other new information. It is also used to express background information. Clitic-VSO order, however, is used in the body of the discourse to convey the main event line and is not associated with new information. This type of distinction between two orders supports Payne's analysis. The arguments for a syntactically underlying order have nothing directly to say about the issue of pragmatic markedness. All they tell us is that both SVO and clitic-VSO are well-motivated syntactically[26]. If D. Payne is correct in her claim that clitic-VSO order is the pragmatically basic word order, and yet the underlying

syntactic order is SVO, as argued here, then an interesting situation arises: underlying order and pragmatically unmarked order may be coexistent, mutually compatible, but nonequivalent in a single language. That is, a language may have, as does Yagua, more than one kind of 'basic word order', depending on whether the focus is pragmatics or syntax. But this means that the notion of 'basic word order', assumed to be a unitary notion in most typological studies, is in fact imprecise and in need of further qualification. It is necessary to state generalizations explicitly in terms of either underlying or pragmatically unmarked order. Since basic word order fails to distinguish these two notions, it itself is of little use and should be avoided as a scientific term.

It is also worth noting that, contrary to Payne 1986, Yagua violates no word order universals under this account (e.g. genitive-head noun order, postpositions, and suffixation which would be difficult to account for if Yagua were VSO). The conclusion here then 'solves' the apparent typological strangeness of Yagua by providing detailed, independently justified arguments that Yagua is not VSO.

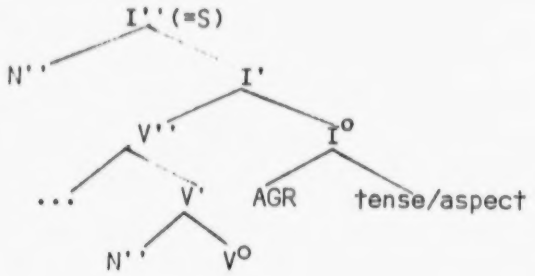
7 Some diachronically oriented speculation

Like most Amazonian languages (Derbyshire 1987), Yagua word order has certain features which seem to indicate that it was historically SOV. For example, it has postpositions, the genitive precedes the NP head, subject agreement morphemes attach to the left edge of their host (V) and tense/aspect morphemes attach on the right edge of the verb.

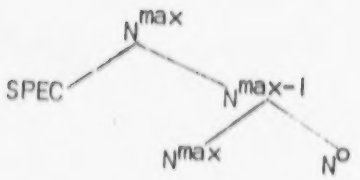
If we assume that Yagua word order was SOV at an earlier period, the above facts are expected for theoretical reasons. Generally, perhaps always, Case in SOV languages is assigned uniformly to the left (cf. Travis 1984). The X'-structure of such languages is as in (85):

(85)

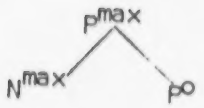
a.



b.



c.



Case will be assigned to the immediate left in all X' projections from X^o to N^{max}, thus accounting for the fact that OV order, postpositions, and genitive-N^o orders are generally found in the same language.

The obvious problem for this is the assignment of nominative Case from I^o to subject, $[N^{max}, I^{max}]$, position. If adjacency is assumed once again as a constraint on Case-assignment, then nominative Case assignment will be routinely blocked in SOV languages, unless the Case-assigner, AGR (tense and aspect are not generally assumed to play a direct role in Case assignment), raises and adjoins to a position right-adjacent to the subject (the subject cannot lower for numerous reasons, e.g., the Empty Category Principle and the Binding Conditions). If AGR is realized by an independent term (i.e., if it is not a clitic or an affix), then it will occur in second position in SOV languages. If, as GB theory maintains, AGR, not tense or aspect, is the relevant Case-assigning portion of I^o, then there is no motivation for raising tense or aspect along with AGR. Assuming that, as a dependent term, tense/aspect must attach to a host then the simplest solution is to allow it to attach leftward to the first host, thus deriving the fact that tense and aspect are verbal suffixes in SOV languages. Now, if AGR is a dependent term, and we assume that I^n and V^n are both [+V,-N], then the logical host for AGR, if it is sensitive to features, is the verb. After assigning nominative Case, then, AGR attaches to the first [+V,-N] morpheme boundary it encounters, immediately predicting it to be a verbal prefix.

This also accounts for the fact that Set I clitics attach to the left, while Set II clitics attach to the right. D. Payne 1985:32,151ff observes that the object often appears at the right periphery of the VP, as we have also seen in various examples above. Let us suppose that the language was indeed SOV at an earlier stage but that it then developed the rule of rightward object movement, transporting the object from preverbal position to the right of the verb (adjoining to V or VP). This would be very much like the synchronic rule of 'Heavy-NP Shift' that is observed in many SOV languages of the Amazon (e.g. Pirahã) and elsewhere. At this juncture, we may hypothesize, [+V] Case began to be assigned to the right, eliminating the need for the rightward movement rule, allowing the object to be generated directly in postverbal position, effectively changing the language to SVO, and perhaps forcing Set II clitics to change from proclitics to enclitics in order to assign Case to the postposed object. If we further suppose that the pressure to maintain unidirectionality of Case assignment forced all Cases to be assigned to the right, we begin to develop an interesting account for the present system, including the prefixal status of Set I clitics (they did not need to become enclitics initially because there was no rule comparable to Object Postposing for other arguments) versus the suffixal status of Set II clitics.

If this reasoning is on the right track, it suggests that the morphological asymmetry noted by typologists, that tense and aspect go on one side and agreement on the other, may turn out to have a syntactic explanation, rather than require the postulation of a distinct rule component to handle the derivations (see Baker 1988 for more suggestions on the morphology-syntax interface).

One other observation should be made. If dependent terms of agreement are not sensitive to features and if they are not [+C], then they will not be prohibited from attaching to their doubles as are Yagua clitics. If a dependent term did attach to its double rather than to the head of its phrase, this would result in the situation described by Nichols 1986 for so-called 'dependent marking' languages. Therefore, there really is no mystery that some languages mark agreement on the head while others mark it on the complement. This would be epiphenomenal, deriving from the theories of Case and syntactic features. Moreover, it is worth investigating, in light of these results, whether or not the criteria proposed by Zwicky and Pullum 1983 for distinguishing between affixes and clitics are derivable.

8 Implications for lexical theories

The analysis proposed above, whereby Yagua clitics attach to their hosts in the syntax, is incompatible with theories which derive inflection in the lexicon, such as Lexical-Functional Grammar (LFG) and Lexical Phonology (LP). The problem is especially acute for LP since Set I clitics undergo otherwise word internal processes which Set II clitics fail to undergo. In LP, this can only be explained if Set I

clitics attach in the lexicon. On the other hand, my account is able to derive these facts straightforwardly without positing lexical strata, by means of independently required principles of Case theory.

While various suggestions come to mind as to how one might treat Yagua clitics in LP or LFG, there are severe problems for any framework that would require Set I clitics to attach in the lexicon. I believe that these problems are in fact insurmountable:

1. VSO versus SVO word order alternations.
2. The derivation of distinctions between Set I and Set II clitics.
3. Index percolation and the derivation of Yagua reflexive facts.

The Case-based analysis derives all of these facts, as we have already seen. Let us review the facts and why they are problematic:

Clitic-VSO order is derived in my account from SVO order by V-to-AGR raising, motivated by the need for the AGR clitic to receive Case. However, if Set I clitics appear on their hosts at D-structure, they will already be morphologically visible (see (34) above) and there will thus be no motivation for the verb to move. Also, SVO order with no clitic on the verb would be syntactically indistinguishable from the ungrammatical order S clitic-VO, since both would have the same structure: S[$_V$(clitic) V]O. The syntax could not tell whether or not the material in parentheses is present. Since neither LFG nor LP can account for this fact, the present account is preferable.

The present analysis derives the Set I versus Set II clitic distinctions as to degree of attachment to the host via the adjacency requirement on Case-assignment (see section four above). A lexical analysis could only stipulate this distinction by placing the two types of clitics in different strata. That this is nothing more than a stipulation is clear: the situation could easily be reversed. Neither LFG nor LP would have any independent reason to prohibit the opposite ordering, attaching Set II clitics earlier than Set I clitics.

If the clitic were part of the host at D-structure, the syntax would have no way to assign an independent index to it. This is because, as argued by Simpson 1983, bracket erasure renders subparts of words opaque to the syntax (also known as the Lexical Integrity Hypothesis). At D-structure, for example, under LFG or LP, sa-púúchu 'S/he carries...' would have the structure in (86):

(86) [$_V$ sapúúchu] 'S/he carries...'

There is no sense, given such a representation, in saying that sa bears an index. In fact, it would be impossible to claim this in either LFG or LP. Therefore, neither LFG nor LP is able to predict the reflexivization facts of Yagua which the present analysis captures under

the independently required device of index-percolation (motivated, again, by the adjacency requirement on Case-assignment).

I conclude, therefore, that clitic-doubling in Yagua presents a serious challenge to theories in which inflectional morphemes are attached to their hosts in the lexicon.

9 Conclusion

I have argued in this paper that the wide variety of facts associated with clitic doubling in Yagua results from Case theory (which includes the notion of morphological visibility), the parametrization of Yagua clitics as [+C,-A], and the notions of inclusion and exclusion. All of these concepts and principles are independently available in UG.

To the degree that this analysis is successful in accounting for the facts, its heavy dependence on Case theory provides evidence in favor of the only syntactic theory which contains a theory of Case, namely, GB. The focus on Case makes comparison with other theories lacking such a subcomponent interesting since it is unlikely that this array of facts could be given such a unified analysis any other way.

NOTES

*Yagua is spoken by approximately three thousand individuals in northeastern Peru. It is the only extant member of the Peba-Yaguan family. Detailed studies of Yagua grammar are D. Payne 1985, T. Payne 1985, and Payne and Payne 1988. I want to thank Tom and Dorrie Payne for their careful work on Yagua, without which the present study would not exist, and for discussions on the analyses and data presented here. Thanks also to audiences at the Working Conference on Amazonian Languages at the University of Oregon, the University of California, San Diego, and the University of Pittsburgh. Sole blame for the analyses and interpretations of the data below must, unfortunately, be attributed to me. This research was funded by NSF Grants BNS 8405996 and BNS 8617854, NEH Grant RX-20870-87, the University of Oregon Foundation, and the Summer Institute of Linguistics. As more articles appear on Amazonian languages, it is hoped that attention will also be drawn to the plight of the speakers of these languages who often stand in imminent danger of losing their traditional lands to Western economic interests.

1. I will not count the clitic in determining adjacency, so that in V-clitic NP, V is adjacent to NP. Reasons for this appear as the discussion progresses.

2. Coreference clitics are omitted here. See section five.

3. Jelinek 1984 claims that doubles in some languages are nonargumental, appositive to the true arguments, the clitics. I will interpret this to mean for those languages that the doubled NPs are in A'-positions.

4. Weak crossover refers to '...structures containing a pronoun P referentially dependent on a wh-trace (or a QR-trace) when wh-movement (or QR) has taken place from some A-position A* to some A'-position A'* such that A'* c-commands both P and A* and neither P nor A* c-command the other' (Sportiche 1985:467).

5. Sells 1984:15 defines a resumptive pronoun as '...a pronoun that is operator bound.'

6. 'Corpus' here refers to all the references to Payne in the bibliography.

7. See DiSciullo and Williams 1987:25ff, for arguments that words may have multiple heads, such that α may be the head of X^0 for feature F_1 and β the head of X^0 for feature F_2.

8. Borer's (1984) complement-matching proposal is another way of expressing the role of clitics in agreement, although she does not state this explicitly.

9. Heads assign a single Case, as expressed in the Uniformity Condition (Chomsky 1986a): 'If α is an inherent Case-marker, then α Case-marks NP if and only if α θ-marks the chain headed by NP'. Although verbs are not strictly inherent Case assigners, it would nevertheless not be possible to allow them to assign Cases to expletives (such as Yagua Set II clitics) and arguments simultaneously. This would render the expletive 'visible' at LF independently of a lexical head. Such an expletive would not be able to enter into a chain with an argument NP which already has Case since a chain may only bear a single

Case. But an expletive which neither enters into a chain nor is interpreted via its relation to a head will at once require an independent interpretation (since it is visible) but will have none (it is not itself an argument). Thus, an expletive assigned Case by a verb in this fashion would violate the principle of Full Interpretation (Chomsky 1986a) and is not allowed. Therefore, verbs may only assign Case to a nominal if it is an argument or in a θ-chain.

10. See Jaeggli 1982, Borer 1984, and Everett 1986 on Spanish clitic doubling, the Case conflict produced, and its resolution.

11. See section six for evidence for a VP node in Yagua.

12. What Safir 1981 really intends is that the licensing of pro be interpreted as epiphenomenal, a result of whether or not a language allows a Case to go unexpressed. Even if his arguments on pro do not go through, the proposal that Case is obligatorily manifested phonetically, except where otherwise parametrized, seems a valid one of direct relevance here.

13. See note 9 on why Vs cannot assign Case to certain expletives.

14. Visibility, (34), does not allow the clitic to attach to its double since the double is not independently visible.

15. A Set II clitic will not be included since this extra degree of attachment would not be motivated. One way to conceive of inclusion is movement-via-substitution whereas being neither included nor excluded is adjunction. I assume that inclusion is only allowed if motivated.

16. My purely syntactic account differs from that of Payne and Payne 1988, which analyzes indirect and direct object interrelations in terms of a notion of definiteness. The reader is referred to that work for an alternative analysis. Since I am able to account for all the facts they record, plus a number of others, without need for formally undefined terms (in their usage) like indefiniteness, I consider the syntactic analysis superior. I suspect that the 'pragmatic' effects they note are epiphenomenal, although space does not allow further discussion of this issue.

17. This unaccusative structure, where the single argument receives structural Case (as shown by the Set II clitic) is at once a problem for 'Burzio's Generalization' (Burzio 1986:185ff), that unaccusative verbs do not assign Case, and Belletti's hypothesis that unaccusatives only assign inherent Case. But there is just too little data on these structures to press the issue here. Note, too, in examples like (52) and (53) that Case may be assigned in equatives even in the absence of an overt verbal element. I have no comment on this except to speculate that Case in equatives results from the relation of predication between the predicate nominal and the unaccusative subject.

18. This indexing is an extension of the complement NP's index. It will not itself bind the NP (which would be a violation of Binding Condition C, Chomsky 1982:6ff) if unlike categories may not bind each other (the 'offending' index is on a V), or alternatively, if only arguments were potential binders. In either situation, the only category which can bind an NP out of the dominating X^{max} is the NP complement, via its index on its dominating X^{max}. The Binding Conditions are:

Binding Conditions:

A. An anaphor must be bound in its minimal governing category (roughly, its minimal containing NP or S).

B. A pronoun must be free in its minimal governing category.

C. A referential expression is free.

Binding: α binds β if α c-commands β and α and β are coindexed.

C(onstituent)-Command:

α c-commands β if and only if the first branching node which dominates α dominates β or α is X^0 and β is in the maximal projection of α.

19. Index percolation is only motivated (and only possible) if the clitic is outside its host at D-structure. See section eight.

20. The c-command domain for Yagua subjects, as in other languages, will of course always be S, whether or not the subject is doubled.

21. Neither possessors nor obliques should be able to bind at D-structure. D. Payne 1985:183ff presents evidence from causatives that subjects may bind at D-structure. This is permitted by my analysis since doubling of subjects, as mentioned in note 20, does not affect the subjects' c-command domain.

22. I have found no clear data on the possibility of c-command within VP, but see example (83).

23. T. Payne 1985:46ff notes that there are examples of <u>jíy-/yí-</u> used in isolated clauses or with antecedents outside their S^{max} in text material. But, since he also notes that these are infrequent and that native speakers tend to edit such clauses so as to place the antecedent in the same clause as the anaphor, I will consider such examples to be performance errors.

24. See Hale 1973 for arguments that clitics may attach to 0-morphemes (or empty categories).

25. Other potential underlying orders are not considered since they do not appear at S-structure and would therefore involve <u>absolute neutralization</u>.

26. The clitic adds extra information (person, number, animacy) which helps identify its double, the true argument. That the clitic thus aids in communication by redundant specification might explain its lower pragmatic markedness. If this is correct, then the syntactic phenomenon of agreement might indeed have direct pragmatic relevance. I will not speculate further on this here, however.

REFERENCES

Aoun, Joseph. 1981. The formal nature of anaphoric relations. Cambridge, MA: MIT dissertation.

Baker, Mark. 1988. Incorporation: a theory of grammatical function changing. Chicago: University of Chicago Press.

Belletti, Adriana. 1982. 'Morphological passive' and pro drop: the impersonal construction in Italian. Journal of Linguistic Research. 2.1-33.

_____. 1988. The Case of unaccusatives. LI. 19.1-34.

Borer, Hagit. 1984. Parametric syntax. Dordrecht: Foris.

_____. ed. 1986. Syntax and semantics 19: the syntax of promotional clitics. Orlando: Academic Press.

Bresnan, Joan. ed. 1982. The mental representation of grammatical relations. Cambridge, MA: MIT Press.

_____, and Sam Mchombo. 1987. Topic, pronoun, and agreement in Chichewa. Lg. 63.741-82.

Burzio, Luigi. 1986. Italian syntax. Dordrecht: D. Reidel.

Chomsky, Noam. 1981. Lectures on government and binding. Dordrecht: Foris.

_____. 1982. Some concepts and consequences of the theory of government and binding. Cambridge, MA: MIT Press.

_____. 1986a. Knowledge of language: its origin, nature, and use. New York: Praeger.

_____. 1986b. Barriers. Cambridge, MA: MIT Press.

_____. 1988. Language and the problems of knowledge: the Manuagua lectures. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1984. A' bound pro vs. variable. MS. Universita de Venezia and MIT.

Chung, Sandra. 1983. The ECP and government in Chamorro. NLLT. 1.207-44.

Davies, William. 1986. Choctaw verb agreement and Universal Grammar. Dordrecht: D. Reidel.

Derbyshire, Desmond. 1987. Morphosyntactic areal characteristics of Amazonian languages. IJAL. 53.311-26.

DiSciullo, Anna Maria and Edwin Williams. 1987. On the definition of word. Cambridge, MA: MIT Press.

Dooley, Robert. 1982. Options in pragmatic structuring of Guarani sentences. Lg. 58.307-31.

Emonds, Joseph. 1985. A unified theory of syntactic categories. Dordrecht: Foris.

Everett, Daniel. 1986. Pirahã clitic doubling and the parametrization of nominal clitics. MIT working papers in linguistics, vol. 8, ed. by Naoki Fukui, Tova Rapoport, and Elizabeth Sagey, 85-127.

_____. 1987. Pirahã clitic doubling. NLLT. 5.245-76.

Fabb, Nigel. 1984. Syntactic affixation. Cambridge, MA: MIT dissertation.

Givón, Talmy. 1976. Topic, pronoun, and grammatical agreement. Subject and topic, ed. by Charles Li. Orlando: Academic Press.

_____. 1984. Syntax: a functional-typological perspective. Amsterdam: John Benjamins.

Hale, Kenneth. 1973. Person marking in Walbiri. A festschrift for Morris Halle, ed. by Stephen Anderson and Paul Kiparsky. New York: Holt, Rinehart, and Winston.

Jaeggli, Osvaldo. 1982. Topics in Romance syntax. Dordrecht: Foris.

Jeanne, Laverne. 1978. Topics in the grammar of Hopi. Cambridge, MA: MIT dissertation.

Jelinek, Eloise. 1984. Case and configurationality. NLLT. 2.1.

Johnson, Kyle. 1987. Against the notion "SUBJECT". LI. 18.354-61.

Kayne, Richard. 1983. Connectedness and binary branching. Dordrecht: Foris.

Kiparsky, Paul. 1985. Some consequences of lexical phonology. Phonology Yearbook. 2.85-138.

Koopman, Hilda. 1984. The syntax of verbs. Dordrecht: Foris.

Lasnik, Howard and Mamoru Saito. 1984. On the nature of proper government. LI. 15.235-89.

Mohanon, K.P. 1986. The theory of lexical phonology. Dordrecht: D. Reidel.

Nichols, Johanna. 1986. Head marking and dependent marking grammar. Lg. 62.56-119.

Payne, Doris. 1985. Aspects of the grammar of Yagua: a typological perspective. Los Angeles: UCLA dissertation.

_____. 1986. Basic constituent order in Yagua clauses: implications for word order universals. Handbook of Amazonian languages, vol. 1, ed. by Desmond Derbyshire and Geoffrey Pullum, 440-68. Berlin: Mouton de Gruyter.

_____. 1987. Morphological characteristics of lowland South Americal languages. MS. University of Oregon.

_____, and Thomas Payne. 1988. Yagua. Handbook of Amazonian languages, vol. 2, ed. by Desmond Derbyshire and Geoffrey Pullum. Berlin: Mouton de Gruyter.

Payne, Thomas. 1983. Yagua object clitics: syntactic and phonological misalignment and another possible source of ergativity. Interplay of phonology, morphology, and syntax, ed. by John Richardson, Mitchell Marks, and Amy Chukerman et al., 173-84. Chicago Linguistics Society.

_____. 1985. Participant coding in Yagua discourse. Los Angeles: UCLA dissertation.

Picallo, Carme. 1984. The Infl node and the null subject parameter. LI. 15.75-102.

Pulleyblank, Douglas. 1986. Tone in lexical phonology. Dordrecht: D. Reidel.

Rizzi, Luigi. 1985. On the status of subject clitics in Romance. Studies in Romance, ed. by Osvaldo Jaeggli and C. Silva-Corvalan. Dordrecht: Foris.

Rothstein, Susan. 1983. The syntactic forms of predication. Cambridge, MA: MIT dissertation.

Safir, Ken. 1981. Inflection-government and inversion. The Linguistic Review. 1.417-67.

_____. 1987. The syntactic projection of lexical thematic structure. NLLT. 5.561.601.

Saxon, Leslie. 1986. Dogrib pronouns. San Diego: UCSD dissertation.

Sells, Peter. 1984. Syntax and semantics of resumptive pronouns. Amherst, MA: University of Massachusetts dissertation.

Simpson, Jane. 1983. Aspects of Walbiri morphology and syntax. Cambridge, MA: MIT dissertation.

Sportiche, Dominique. 1985. Remarks on crossover. LI. 16.460-9.

Sproat, Richard. 1985. On deriving the lexicon. Cambridge, MA: MIT dissertation.

Stowell, Timothy. 1981. Origins of phrase structure. Cambridge, MA: MIT dissertation.

Travis, Lisa. 1984. Parameters and effects of word order variation. Cambridge, MA: MIT dissertation.

Williams, Edwin. 1980. Predication. LI. 11.203-38.

Zwicky, Arnold and Geoffrey Pullum. 1983. Cliticization and inflection: English _n't_. Lg. 59.502-13.

TABARU PHONOLOGY AND MORPHOLOGY

Edward A. Kotynski

CHAPTER 1 INTRODUCTION
CHAPTER 2 STRESS
1 Concepts of stress
1.1 Structuralist phonology
1.2 Generative phonology
1.3 Metrical phonology
1.3.1 Parameters
1.3.1.1 Foot structure
1.3.1.2 Word structure
1.3.1.3 Extrametricality
1.3.2 Assumptions
2 Presentation of facts
2.1 Monosyllabic words
2.2 Polysyllabic words
2.3 Stress patterns
3 Presentation of accounts
CHAPTER 3 MORPHOLOGY
1 Verbal morphology
1.1 Prefixes
1.1.1 Pronominal subject prefix (PS-)
1.1.2 Pronominal object prefix (PO-)
1.1.3 Reduplication
1.1.4 Applicative 1 (dV-1)
1.1.5 Applicative 2 (si-)
1.1.6 Reciprocal (ka-)
1.1.7 Surrounding (sa-)
1.1.8 Many subjects, few objects (ri-)
1.1.9 Continuative (ta-)
1.1.10 Habitual (taba-)
1.1.11 Applicative 3 (g-)
1.2 Suffixes
1.2.1 Locational suffixes (-LOC)
1.2.2 Negation (-uwa)
1.2.3 Reduplication, emphatic negation (RDP-)
1.2.4 Past tense (-oka)
1.2.5 Already (ou)
1.2.6 Repetitive (-oli)
1.2.7 Continuative (-osi)
2 Nominal morphology
2.1 Reduplication
2.2 Locational suffixes
3 Derivational morphology
3.1 Verb derivations
3.2 Noun derivations

CHAPTER 4 PHONOLOGY
1 Inventory of segments
  1.1 Consonants
  1.2 Vowels
  1.3 Semivowels
2 Syllable structure
3 Affixation
  3.1 Rules associated with the g- prefix
  3.2 Reduplication
  3.3 Vowel deletion
  3.4 vowel epenthesis
  3.5 The dV- prefix
  3.6 Stress
  3.7 Pronominal prefixes
4 Lexical phonology
5 Lexical phonology of Tabaru
6 Derivations
APPENDICES
A List of Abbreviations
B Interlinear Tabaru text
REFERENCES

Introduction

In this paper we examine the Tabaru language in light of recent phonological theory.[1]

First, we review the history of stress within various theories and discuss both strengths and weaknesses. After concluding that previous theories are not sufficiently predictive or restrictive in terms of general theory, we discuss metrical phonology (Harris 1983, Mohanan 1986) and apply that theory to the Tabaru language.

Second, after reviewing the morphology of Tabaru, we discuss the notion of syllable and its structure within the theory of CV phonology (Clemonts and Keyser 1983). We then analyze the Tabaru syllable within this framework, and show that the Resyllabification Convention as stated in Clements and Keyser (1983) is crucial for an adequate analysis of Tabaru.

Third, we review briefly the theory of lexical phonology (Mohanan 1986) and present an analysis of Tabaru using lexical phonology.

The above theories used together provide a foundation for the analysis of phenomena difficult within the framework of generative phonology alone. Although generative theory is adequate to provide an analysis for stress in Tabaru, the theory is not restrictive enough. It makes no restrictions on the patterns of stress which can occur in human language. Metrical phonology provides a more restrictive theory. In analyzing Tabaru using generative phonology it appears that there is a contrast between [u] and [w] and between [i] and [y]. The theory of CV phonology provides a criterion upon which to predict the occurrence of [w] and [y] from underlying [u] and [i]. Lexical phonology provides a coherent analysis of the phonology between the morphemes by allowing phonological processes to be interspersed between morphological processes, which is not normally allowed within generative phonology.

The Tabaru language is spoken on the island of Halmahera in the province of Maluku, Indonesia. It is a West Papuan Phylum language. Some of the languages of this group are spoken in the northern peninsula of Halmahera and nearby islands, the majority being spoken in the Bird's Head of Irian Jaya (Voorhoeve 1981). In the late 19th and the early 20th centuries, Dutch missionaries and linguists were interested in the North Halmaheran languages because of how much they differed from the languages in the rest of Halmahera and because of the regularity of the sound correspondences between them. Van der Veen (1915) wrote a dissertation demonstrating the differences between the North Halmaheran languages and the Austronesian languages of the area. He showed beyond reasonable doubt that the North Halmaheran languages are significantly different, and that they are not Austronesian languages. In more recent times, further evidence has come to light (Voorhoeve 1984) to demonstrate the relationships between the languages of North Halmahera and the Bird's Head languages of Irian Jaya. There are approximately ten to fifteen thousand speakers of Tabaru, with two known major dialects, Tabaru-nyeku (the highland Tabaru) and the Tabaru-adu (the lowland Tabaru).

NOTES

I express here my appreciation to the members of my Advisory committee: Dr. John Crawford, Dr. Stephen Marlett and Dr. David Marshall. I thank especially Dr. Marlett for the interest, criticism and direction he gave as this thesis was being written.

I am indebted to the University of Pattimura in Ambon, Indonesia, under whose sponsorship my field work with the Summer Institute of Linguistics was conducted.

Agustina Tatuh Nijhatiti of Mangga Dua, Ternate who helped us in language learning and Subaduano of Suli, Ambon, for assistance in glossing texts, deserve my greatest thanks for their sharing of the Tabaru language. And my heartfelt thanks to Mrs. Gerrie Natelborg for her invaluable help in rendering Dutch materials into English.

[1]The Tabaru language has been referred to in the literature as both Tabaru and Tobaru. I follow Fortgens here in using Tabaru. He gives as his reason (Fortgens 1928:390) that the people call themselves Tabaru, although people from Ternate and the other related languages call them Tobaru.

This paper uses examples from my field notes and Fortgens' (1928) grammar notes and folk stories.

CHAPTER 2
STRESS

1 Concepts of stress

Stress has been dealt with in various ways in the past by both structuralist and standard generative linguists. An examination of these theoretical approaches to stress, including a discussion of their strengths and limitations along with an introduction to metrical phonology will be the basis used to argue for an analysis of Tabaru stress.

1.1 Structuralist phonology (phonemics)

The notion that stress can be phonemic comes from the structuralist approach to phonology. Stress is considered a phoneme if, for example, there are minimal pairs in a language which contrast only in stress. Bloomfield (1933) in his discussion on the phoneme goes a step further by making a distinction between primary and secondary phonemes. Another way to differentiate between them might be to use the term phoneme to mean primary phonemes as in the following minimal pair:

bit
pit

and equate suprasegmentals with secondary phonemes, such as stress or pitch:

'contest	noun
con'test	verb
at four o'clock?	question
at four o'clock.	answer

(Bloomfield 1933:91). He further claims that the place of stress can be a primary phoneme if the stress on monomorphemic forms varies and the meaning is dependent upon it. He cites Russian (Bloomfield 1933:111):

'gorot	'city'
mo'ros	'frost'

Pike (1947:76-77) holds to a similar view although not making the distinction between primary and secondary phonemes. Using as an example the hypothetical language Kalaba he demonstrates phonemic stress as follows:

'tomas	'black'
ta'mof	'blue'

With this evidence he says that in Kalaba stressed and stressless syllables contrast in analogous environment. Stress is phonemic; that is, stress is a phoneme in this language. However he also points out that lack of stress is not a phoneme.

As late as 1961, Gleason (1961:40ff) suggested that minimal pairs be sought in a language for the purpose of determining if stress is phonemic. He uses English data in his arguments:

'pervert	noun
per'vert	verb

He then goes on to discuss three distinct stress phonemes in English.

All of the above approaches are essentially the same; they state that stress can be a phoneme in some languages. Only Bloomfield makes the distinction between primary and secondary phonemes, thus allowing for a qualitative difference between the stress placement in English noun and verb forms 'pervert, per'vert and the Russian examples where stress placement is crucial in determining the meaning of the word, not grammatical category. Although in this view it is possible to make language-specific generalizations about stress which are stated in prose form, those generalizations are not specific enough to establish formal rules and thus would not eliminate the need to mark stress in the phonemic representation. For example, in Tabaru there is contrastive evidence for stress in such words as:

'caana	'thousand'
ca'ana	'pants'

A structuralist might look at Tabaru and observe that stress in Tabaru usually falls on the penultimate syllable and yet fail to capture this generalization by positing a stress phoneme and indicating stress in the phonemic representation. We can see that although structuralists try to generalize as much as possible, the theory can predict stress only when it occurs without exception on the phonetic level.

1.2 Generative phonology

In structuralist phonology, there are three levels of phonological representation: the phonetic, the morphophonemic, and the phonemic level. Generative phonology accepts only two levels: the surface form (phonetic) and the underlying form. Generative phonologists completely rejected the phonemic level of representation and in doing so eliminated the possibility of phonemic stress. This barred them from allowing stress to be represented in the lexicon (where phonemic stress is placed in structuralist theory) and tied stress instead to phonological rules designed to predict stress (Chomsky & Halle 1968:145ff). This view was a quantum leap from phonemics in which stress was largely unpredictable. In the generative view, stress is applied to segments by means of phonological rules which place the feature [+stress] on the particular feature matrix (segment) to which stress applies. The rule below is an example of such a rule which might be posited for Tabaru.

$$V \rightarrow [+\text{stress}] \; / \; __C_oVC_o\#\#$$

However, the feature [+stress] is different from the other distinctive features in two ways. First, all the other features are binary. Stress is not necessarily binary. To handle stress in English, for example, it is necessary to posit various intensities of stress:

$$[2\ \text{stress}] \rightarrow [3\ \text{stress}] \; / \; _____C_o[1\ \text{stress}]$$

(Chomsky and Halle 1968:116), thus making it a non-binary system. Second, because the theory disallows stress in the lexicon, it never occurs in the underlying form as can the other distinctive features.

While this theory claims that stress is generally not idiosyncratic, its formalisms are too powerful. The model rejects the idea of idiosyncratic stress in favor of generalizations but provides no constraint upon which the generalizations should be based. As a result, although it predicts that stress will follow certain patterns, it also predicts patterns that do not actually occur in human language. Thus the rules that can be written may allow for stress to occur anywhere, when in fact there are constraints as to where stress can occur in human language. A further development was needed to restrict the theory, giving it only the power to formalize (make predictions) about phenomena which actually occur in human language. For

example, in standard generative phonology the following rule is possible:

V --> [+stress] / C_CVCVCVCV##

But this rule, which is allowed by the formalism, is not well motivated by the facts in any known language.

Another shortcoming in standard generative theory is that it allows for no notion of the syllable. This is because it views phonology as a linear string of segments and boundaries to which rules are applied in order to produce the surface form. The linear string posited includes neither the syllable nor anything else above the segmental level. This leads to problems, because in proposing the structure of tone and stress, it is desirable to make reference to more than one segment at a time. Standard generative theory does not recognize the higher levels and as a result does not adequately handle tone and stress.

In the middle 197Øs, studies began to focus on suprasegmentals such as stress (Liberman and Prince 1977) and tone (Goldsmith 1976). It soon became clear that it was necessary to posit a hierarchical structure of phonological representation so that notions such as tone and stress could be handled adequately. The theories of autosegmental phonology and metrical phonology have been developed as a result of these studies. An analogy of the hierarchy common to these theories is found in Halle and Vergnaud (1982:65):

> The major insight lying at the base of autosegmental phonology is that the phonological representation is composed not of a single sequence of entities roughly resembling a line of type, but rather that the phonological representation is made up of several parallel sequences of entities, resembling thus more a score for a musical ensemble than a single line of type... What has been novel in autosegmental phonology is that the tones of an utterance are not viewed as diacritics of vowels or syllables; rather the tones are viewed as constituting an autonomous sequence of entities, separate from and equal to the sequence of consonants and vowels that make up what we shall call here the phonemic core of the utterance.

A parallel statement can be made about metrical phonology: syllables, feet and words are also viewed as "autonomous sequences of entities, separate but equal to the sequence of consonants and vowels."

1.3 Metrical phonology

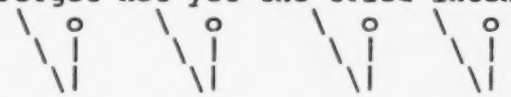
Metrical phonology has its roots in a study of stress which attempts to correct deficiencies in generative theory. On the background to metrical phonology, Hammond (1986:188) says:

> Metrical theory uses some concepts of classical poetic meter to describe stress systems. The particular notion drawn from poetic meter is the metrical foot, a prosodic constituent taken as the building block of poetic verse.

Two concepts of meter exploited by metrical phonology are the iambic and trochaic feet, illustrated in the following (Allison 1975):

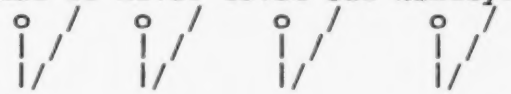
iambic feet

Forget not yet the tried intent.



trochaic feet

Had we never loved sae kindly.



Both of the above lines consist of four binary feet. In the iambic line, the right node of the foot is strong and is designated by a circle and a vertical line. The weak node is represented by a diagonal line alone. In the trochaic line, the left node of the foot is strong and the right node weak. Stress patterns in metrical phonology use these types of feet; the iambic feet are right-headed and the trochaic feet are left-headed.

In this paper we will be following Hammond's (1986) view of metrical phonology. There are various competing views within metrical phonology; however, there are basic

assumptions which are common to all. Before considering these basic assumptions, we discuss the major parameters of metrical phonology below.

1.3.1 Parameters

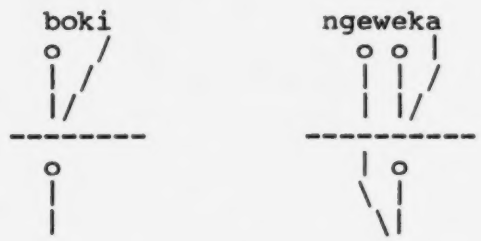
1.3.1.1 Foot structure. There are two parameters to consider in foot structure. The first, foot directionality, defines from which side of the word the trees are built. The structure of feet is represented by tree diagrams which are mapped onto the syllables of a word either from the right or left edge. If stress is counted from the left side of the word then trees are built starting at the left edge of the word and the language has left-branching feet. If stress is counted from the right side of the word then trees are built from the right edge of the word and the language has right-branching feet. In Tabaru, for example, the feet are right-branching and left-headed:1

But languages like Yidin^{y} (see Hayes 1982) are left-branching and right-headed as in the following examples:

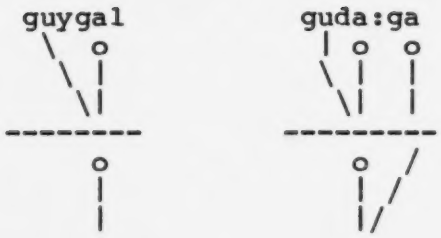
The second parameter, foot headedness, indicates the location of the strong node of the foot. Foot headedness is determined by the distance of stress from the edge of the word. The Tabaru words above are examples of left-headed feet and the examples in Yidin^{y} of right-headed feet.

1.3.1.2 Word structure. Word headedness indicates location of the strong node of the word. The strong node of the word can occur either on the right or left edge of the word. The word tree has no limit to its number of branches, it is "n-ary branching". The conjunction of the strong node of the word and the strong node of a syllable of a foot indicates the place of primary stress of a word. Again we use Tabaru and Yidin^{y} examples:

Tabaru:



Yidin^{y}:

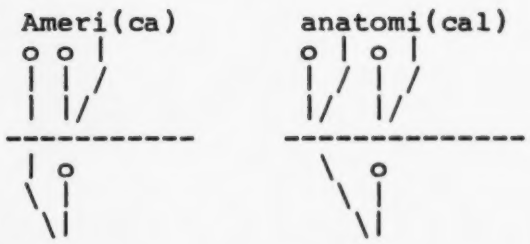


The Tabaru data exemplifies right-headed words and the Yidin^{y} data left-headed words.

1.3.1.3 Extrametricality. There are phenomena in some languages like Nyungan (Hansen and Hansen 1969) in which the ultima is never stressed, or Spanish (Harris 1983) in which a syllable is invisible to poetic meter. The theory of metrical phonology has a formal way to account for these kinds of irregularities, the notion of extrametricality. This notion allows for a syllable at the edge of a word to be considered extrametrical for the purpose of building the foot structure of a word. It is a device which is used to mark syllables which are invisible to rule application. In some languages like Spanish (Harris 1983), the syllable is marked as extrametrical in the lexicon. In other languages like English (Hammond 1986), there are rules which indicate where extrametrical syllables occur:

> The final syllable is extrametrical in a polysyllabic word.

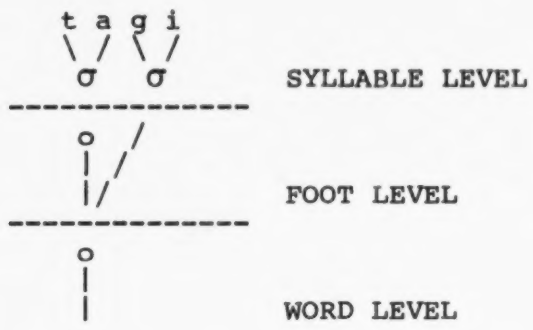
This rule is shown by the following examples (the extrametrical syllable being in parenthesis):



The two options allow for both regular (English) and idiosyncratic (Spanish) exceptions from the tree structure building rules.

1.3.2 Assumptions. Having discussed the major parameters of metrical phonology, we can begin to understand the assumptions which are common to the various views of metrical phonology.

First, within the hierarchical representation, the segments are grouped into syllables, feet and words. The following diagram shows this hierarchy with its three levels.



Syllable structure is discussed in the phonology section.

Second, stress is represented by a hierarchical structure rather than as part of a feature matrix. For example, suppose we have the following rule in generative phonology:

$$V \rightarrow [+stress] \; / \; __C_oVC_o\#\#$$

We then will apply the rule to the Tabaru form:

boki 'cat'

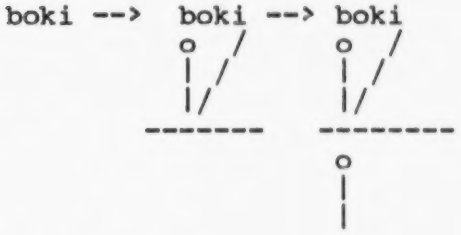
The derivation is:

boki --> 'boki

The equivalent rules in the metrical model are:

1. Build right-branching, left-headed feet.

2. Build right-headed words.



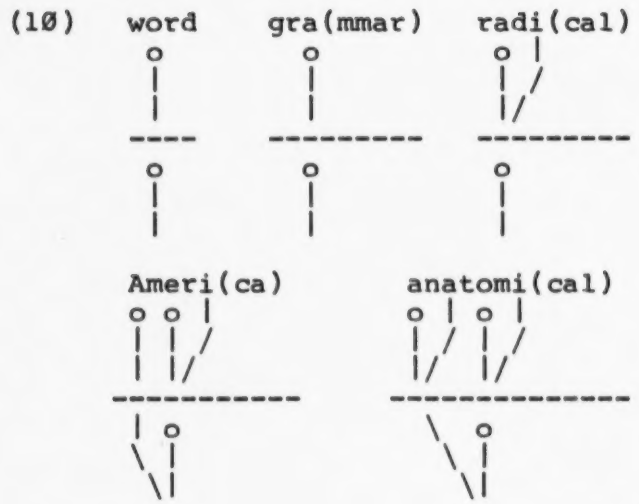
The line with the circle represents a strong node and the line without the circle represents a weak node. The primary stress is determined by following the strong node up from the word (lower) level to the foot (higher) level. The conjunction of the strong word and strong foot nodes indicates primary stress. The result of this above example indicates primary stress on the penultimate syllable.

Third, there is a notion of extrametricality in which a syllable at the edge of a word may be considered extrametrical for the purpose of building the metrical structure. Harris (1983) effectively uses this notion of extrametricality to eliminate the need for lexical diacritics and exception-marking devices in the description of Spanish stress. He gives evidence from poetry and rhyming words as further evidence for the existence of extrametrical syllables. Hammond (1986:190) uses the following rules to demonstrate the notion of extrametricality as it applies to English:

(9)a. The final syllable is extrametrical in a polysyllabic word.

b. Left-headed feet are built right to left.

c. Build a right-headed word tree.



2 Presentation of facts

In order to describe stress in Tabaru, we will present data and propose a number of analyses. First, let us look at the facts about stress in Tabaru. Stems can be prefixed, suffixed and reduplicated. Only suffixation can vary the stress pattern, and regardless of the affixation, primary stress occurs only on the antepenultimate or penultimate syllable. Let us now turn our attention to Tabaru words and the phenomena occurring in them.

2.1 Monosyllabic words

Tabaru has a constraint that noun and verb roots must be composed of polysyllabic morphemes. Monosyllabic words are a restricted set composed of conjunctions, particles and pro-forms (figure 1). They are irrelevant to this paper, and will not be discussed further.

Monosyllabic words:

i	'my'	la	'so that'
so	'so'	de	'and'
ngo	feminine article	o	article
wi	'his'	mi	'her'
ma	'its'	ge	'that'

FIGURE 1

2.2 Polysyllabic words

In figures 2 and 3 we have lists of words, all of which are stressed on the penultimate syllable, the only difference between the lists being that the examples in figure 2 are two syllable words and those in figure 3 are three syllables.

Two syllable:

'goa	'buttocks'	'rio	'footprint'
'mao	'feel'	'rebe	'bet'
'awe	'thread'	'ari	'weep'
'uru	'mouth'	'bira	'rice'
'sira	'first'	'geri	'brother-in-law'

FIGURE 2

Three syllable penultimate stress:

pa'nyake	'illness'	sa'sawi	'coconut shell'
be'laka	'shoulder'	ma'nyanyi	'incense'
sa'maka	'watermelon'	ca'ana	'pants'
ka'kawo	'ash'	do'wongi	'sand'
mo'didi	'two'	mo'doka	'marry'

FIGURE 3

There are, however, also three syllable words having antepenultimate stress as in the following (figure 4):

Three syllable antepenultimate stress:

'odomo	'eat'	'sauku	'warm'
'sowoko	'fruit'	'naoko	'fish'
'akere	'water'	'obiri	'night'
'wekata	'wife'	'koana	'prince'
'gogere	'sit'	'osisi	'urinate'

FIGURE 4

There are monomorphemic words in Tabaru having more than three syllables (figure 5), but they are more difficult to find. Most of these words are stressed on the antepenultimate syllable.

Words having more than three syllables:

go'somanga	'crocodile'	tumu'diingi	'seven'
bo'osuku	'grave'	mogi'ooko	'ten'
me'leunu	'forest chicken'	ga'uwewe	'ancestors'
gu'mutuku	'rope'	kai'lupa	'kepok'
ngu'usumu	'kind of grass'	kasi'naa	'mirror'

FIGURE 5

Stress occurs on the penultimate syllable in words that have suffixes. This is true whether the word lacking a suffix would have penultimate stress, as in figure 6, or antepenultimate stress, as in figure 7.

ROOT:		ROOT + SUFFIX:	
'kula	'give'	ku'lasi	'still give'
'make	'see'	ma'kewa	'not see'
'tike	'look for'	ti'kesi	'still look for'
so'fala	'diligent'	sofa'lawa	'not diligent'
bo'dito	'have misfortune'	bodi'towa	'not have misfortune'

FIGURE 6

ROOT:		ROOT + SUFFIX:	
'okere	'drink'	oke'rosi	'still drink'
'pidili	'loosen'	pidi'luwa	'not loosen'
'odomo	'eat'	odo'muwa	'not eat'
pa'asana	'beat	paasa'nuwa	'not beat'
'punusu	'be full'	punu'suwa	'not be full'

FIGURE 7

2.3 Stress patterns

The following stress patterns appear:

'σσ σ'σσ σσ'σσ σσσ'σσ

'σσσ σ'σσσ σσ'σσσ

It is clear that only penultimate and antepenultimate stress occurs in Tabaru regardless of the length of words. The argument against left-branching feet in Tabaru is that there would be more patterns to account for if this were posited. Counting from the right edge of the word, there are only two stress patterns needed: antepenultimate and penultimate. Counting from the left edge of the word, we would have to

account for four variations in stress pattern, because stress can occur on any syllable from the left edge of the word up to the fourth syllable. We will discuss the simpler explanation.

3 Presentation of accounts

Let us first hypothesize and examine rules to account for the clear cases of Tabaru stress to provide a foundation for examining the less clear cases. Using the above examples, we can propose the following rules to account for penultimate stress:

Account I

Extrametricality rule:
Mark the rightmost syllable extrametrical.

Foot building rule:
Construct right-branching right-headed feet.

Word building rule:
Construct right-headed words.

The derivation follows:

underlying form:

bira 'rice'	samaka 'watermelon'

extrametricallity rule:

bi(ra)	sama(ka)

foot building rule:

word building rule:

As an alternative, let us examine the following account:

Account II

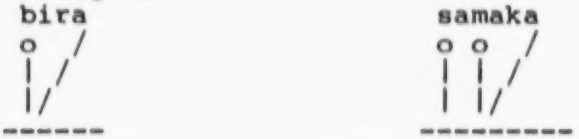
Foot building rule:
Construct right-branching, left-headed feet.

Word building rule:
Construct right-headed words.

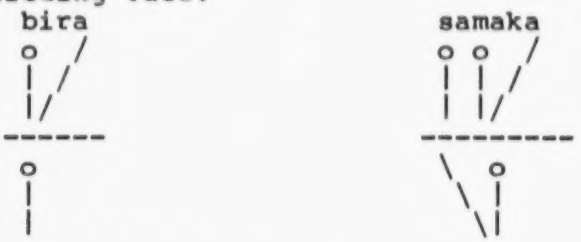
The derivation of Account II is:

underlying form:
bira 'rice' samaka 'watermelon'

foot building rule:



Word building rule:



Both of the above accounts correctly predict primary stress for the above data. Now let us examine data with antepenultimate stress and examine our accounts to determine which will yield the best analysis for these data.

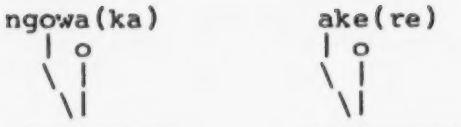
ngowaka 'child' akere 'water'

Account I

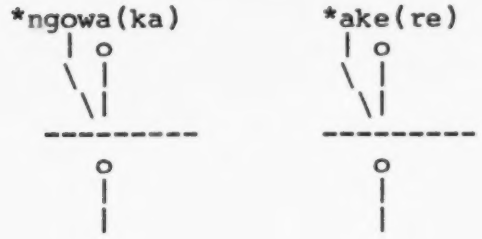
Extrametricality rule:
Mark the rightmost syllable extrametrical.

ngowa(ka) ake(re)

Foot building rule:
Construct right-branching right-headed feet.



Word building rule:
Construct right-headed words.



In light of these new data, Account I produces incorrect stress placement. There is no simple way to salvage the rules because the new data cannot be accounted for without proposing that the final syllable of only some words is marked extrametrical in the lexicon. This unnecessarily complicates the structure by further requiring the marking of the next-to-last syllable as extrametrical as follows:

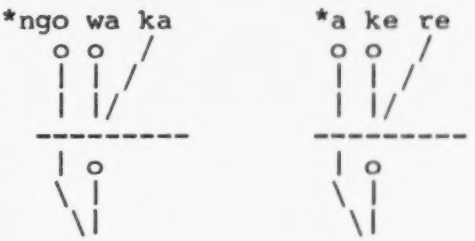
lexical form:	extrametricality rule:	foot construction rule:
ngowa(ka) ->	ngo(wa)(ka) ->	ngo(uwa)(ka)

It is not clear that metrical theory allows for two adjacent extrametrical syllables. We reject this account because it is questionable within the framework of metrical phonology and because the analysis is not well motivated.

Let us now examine Account II. It also does not give correct results as it is currently written. Let us examine the application of the rule to these forms.

Foot building rule:
Construct right-branching, left-headed feet.

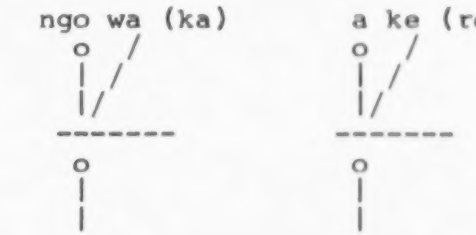
Word building rule:
Construct right-headed words.



The second account cannot be used for these forms without incorrect results as seen above. The prediction here is that the stress is on the penultimate syllable. This account, however, can be revised to produce the correct results. It is necessary only to assume that in the lexicon there are some words which are marked as having an extrametrical syllable. If we make this assumption, then we have words in the lexicon like the following (parenthesis indicating the extrametrical syllable):

ngowa(ka) 'child' ake(re) 'water'

Assuming a lexicon as above, it is necessary only to apply the rules from Account II and we have a correct prediction of stress. The following are the results of the rule application to these forms:



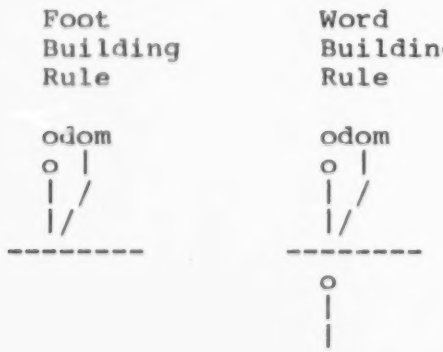
The output of the rules in Account II is now correct because, with the last syllable being invisible to the rules, the stress is placed on the antepenultimate syllable.

Let us now examine the data with suffixes and apply the rules of Account II. The examples in Figures 6 and 7 above show that whether the stress on unsuffixed words is penultimate or antepenultimate, the suffixed forms always take penultimate stress. It is also clear as illustrated in the following forms that the final vowel always appears to drop out when a suffix is added to a word which is stressed on the antepenultimate syllable, as seen in the following examples (figure 8).

'pagara	'lie on belly'	pagar'uwa	'not lie on belly'
'resene	'detour'	rese'nuwa	'not detour'
'pidili	'loosen'	pidi'luwa	'not loosen'
'odomo	'eat'	odo'muwa	'not eat'
'punusu	'be full'	punu'suwa	'not be full'

FIGURE 8

If we maintain that the final syllable in the above words is simply marked extrametrical in the lexicon, it will be necessary to write a rule in order to delete the final root vowel in the event of suffixation. There is an alternative, however, if we again modify our view of these words within the Tabaru lexicon. If we assume that words such as [akere] are represented as /aker/ in the lexicon, this would eliminate our need to delete any vowels when adding suffixes to words which take antepenultimate stress. This also means that we do not need to mark some syllables as being extrametrical (because they don't exist in the lexicon). The underlying representation for [odomo] within this view is /odom/. The surface form is derived as follows:

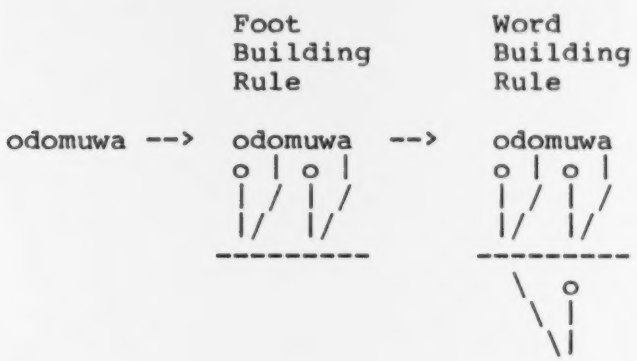


The above rule application places the stress on the first syllable of /odom/. A rule of vowel epenthesis then places the final vowel to produce ['odomo] with its proper stress.

Affixation then creates the following form:

odom+uwa --> odomuwa

The stress rule (Account II) correctly places stress on these forms:



In the case of Tabaru it is not necessary to use the principle of extrametricality in order to assign stress. Here we assume that this is the correct view of the above words in the lexicon. We will examine further motivation for the claim in Chapter 4 where we will also discuss the epenthesis of the final vowel in words stressed on the antepenultimate syllable. The two rules in Account II,

1. Construct right-branching, left-headed feet.
2. Construct right-headed words.

account for primary stress in the Tabaru language.

NOTES

[1]There are other logical possibilities for combinations of branching and headedness which are not exemplified here. There are languages with right-branching and right-headed as well as left-branching and left-headed. The importance of these differing metrical structures becomes apparent on long words which show from which edge of the word the stress is counted. Some languages will not assign metrical structure from the edge of the word as the languages in these examples do, but rather assign structure from the edge of the root.

CHAPTER 3
MORPHOLOGY

The purpose of this section is to examine briefly the kinds of affixation that occur on Tabaru words and to discuss their ordering both in relation to the stem on which they occur and in relation to other affixes. Tabaru has two major lexical classes, noun and verb. There is no adjective class in Tabaru; adjectival predicates belong to the verb class. Because the major word classes in Tabaru are noun and verb, we will examine only the verbal and nominal morphology.

1 Verbal morphology

1.1 Prefixes

The following template indicates the relative order of the Tabaru verbal prefixes, and co-occurrence restrictions. The horizontal lines in the template separate affixes which cannot co-occur.

VERBAL PREFIX TEMPLATE

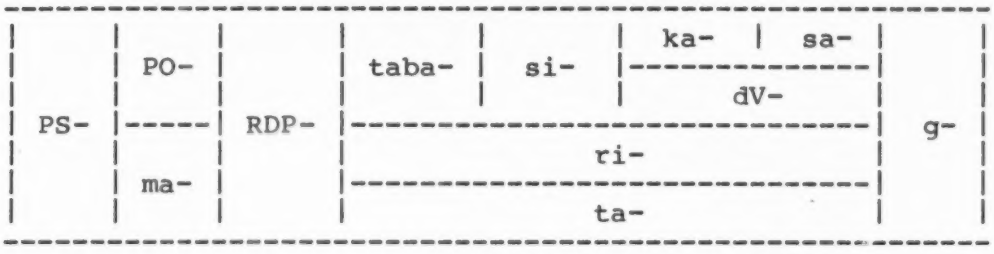


FIGURE 9

A brief description of these Tabaru prefixes (in angle brackets) with examples illustrating their use follows.

1.1.1 Pronominal subject prefix (PS-). These prefixes occur on some intransitive and all transitive verbs.

(1) <no>- ijakasi asa <po>-tagi
<you>-buy.first then <we>-go
'You shop first, then we will go.'

(2) o Tabadiku <wo>-ngose...
ART Tabadiku <he>-say...
'Tabadiku said...'

1.1.2 Pronominal object prefix (PO-). These occur on transitive verbs and as subject markers on some intransitives.

(3) o Tabadiku ma wekata mo-<ki>-siduuru
ART Tabadiku POS wife she-<them>-walk.behind
'Tabadiku's wife followed behind them.'

(4) ma koana wo-<ni>-asoko
ART king he-<you>-call
'The king calls you.'

Tabaru verbs take pronominal prefixes in all contexts. Figures 10 & 11 show their underlying forms:

PRONOMINAL VERB PREFIXES

	Reflex	Unerg	Trans Subjct	Trans Objct	Temp State
1s	to-	to-	to-	i-	i-
2s	no-	no-	no-	ni-	ni-
3ms	uo-	uo-	uo-	ui-	ui-
3fs	mo-	mo-	mo-	mi-	mi-
1pe		mio-	mio-	mi-	mi-
1pi		po-	po-	na-	ina-
1pp	mi-				
1pa	mio-				
2p		nio-	nio-	ini-	ni-
2pp	ni-				
2pa	nio-				
3p		io-			ioki-
3pp	i-		Ø	Øki	
3pa	io-		io-	aki-	
3n		i-			ia-

FIGURE 10

TRANSITIVE PRONOMINAL PREFIXES

SUBJECT \ OBJECT	1s	2s	3m	3f	1pi	1pe	2p	3n(pa)	3pp
1s		toni	toui	tomi			tini	ta(ki)	toki
2s	noi		uoni	moni		mini		na(ki)	noki
3m	uoi	uoni	ui	uomi	uona	uomi	uini	ua(ki)	uoki
3f	moi	moni	uomi	mi	mona	momi	mini	ma(ki)	moki
1pi			poui	pomi				pa(ki)	poki
1pe		mini	miui	mimi			mini	mia(ki)	mioki
2p	nii		uini	mini		mini		nia(ki)	nioki
3pa	ioi	ioni	ioui	iomi	iona	iomi	ioni	ia(ki)	ioki
3pp	i	ni	ui	mi	ina	mi	ni		

FIGURE 11

The use of pronominal prefixes with verbs not in the adjectival class is quite straightforward. However, the use of the pronominal prefixes in the adjectival class depends upon which of two categories of verb they occur with[1]. First, the adjectival verbs which take the unergative prefixes (which correspond closely to the transitive subject prefixes) have an attributive meaning, exemplified in the list below.

Attribute:

solimongo	'stupid'	pululunu	'fat'
cawaro	'smart'	tingoono	'small'
pepeke	'dirty'	tero	'nice'
tipoko	'short'	todomo	'sharp'
bangoro	'beautiful'	pelelaka	'thin'

Second, adjectival verbs, which take the object prefixes, are usually words describing non-volitional actions or temporary states as seen in the following list.

Temporary State:

tuuduku	'depressed'	eto	'drunk'
tokokana	'scared'	tirine	'trembling'
sawini	'hungry'	pelosoko	'tired'
tuunisi	'constipated'	kiwounu	'bored'
surugogo	'choking'	maeke	'ashamed'

The distinction between temporary state and attribute is signaled morphologically by the two verb classes taking different pronominal prefixes. Temporary state verbs take what in the transitive paradigm are the object prefixes (see figure 10 above). This is shown in examples (5) and (6).

(5) iduanga una ka <wi>-punusuali
then he only <him>-satisfied.PER.REP
'Then he was satisfied.'

(6) una <wi>-gogama de <wi>-tirene
he <him>-fever and <him>-tremble
'He is feverish and he trembles.'

The verbs in the Attribute class, as exemplified in (7) and (8), take the pronominal subject prefixes, just like the nonadjectival intransitive verbs (see figure 10 above).

(7) ka <nio>-tingoonosi
only <you.pl>-small.CON
'You continue to be small.'

(8) o Raja Moloko wi namoro <i>-cawaro
ART king Moloko his chicken <it>-smart
'The king of Moloko's chicken is smart.'

1.1.3 Reduplication. The reduplication prefix indicates repetitive action when it occurs on verbs of action, and emphasis when it occurs on adjectival verbs. Both syllables of a two syllable root are reproduced. In all other cases, only one syllable is reproduced. See Chapter 4 for further explanation.

(9) i- <aso>-asoko, Taoko
it-<RDP>-calls, I.it.go.to.sea.
'He repeatedly said, I'm going to the sea'

(10) wo-<rari>-rari
he-<RDP>-cut.grass
'He kept cutting grass.'

1.1.4 Applicative 1 (dV-). The affix dV- indicates the presence of an additional argument of a verb. It can indicate the addition of a benefactive (11) or an indirect object (12).

(11) o ngoi neena o inomo ma oa toma-<da>-make
EXC I this ART food POS good I.REF-<AP1>-meet
'Wow, I've found some good food for myself!'

(12) nako o balusu mini-<do>-ngose nioisene
if ART elders they.you-<AP1>-say you.listen
'When your elders talk to you, you listen!'

1.1.5 Applicative 2 (si-). The addition of si- indicates an additional argument on the verbs as in (13, 14).

(13) o dueku io-<si>-odomo
ART plate they-<AP2>-eat
'They will eat from plates.'

(14) neena o esa de o dea mini-<si>-ngose...
this ART mother and ART father they.you-<AP2>-say
'This is what your mother and father say to you...'

1.1.6 Reciprocal (ka-). The reciprocal morpheme ka- usually occurs with the reflexive prefix ma- (15, 16).

(15) ona yaruangere ima-<ka>-manai
they they.three they.REF-<REC>-friend
'They three are friends with each another.'

(16) o kaso de o titi geena ima-<ka>-sobati
ART dog and ART pig that they.REF-<REC>-friend
'The dog and the pig were friends with one another.'

1.1.7 Surrounding (sa-). This affix meaning, 'surrounding', is rarely found in Tabaru text[2].

(17) o bongana ma bounu wa-<sa>-ame
ART forest POS air he.it-<SU>-smell
'He smelled the air of the woods (surrounding him).'

(18) o ubutu wa-<sa>-rari
ART weeds he.it-<SU>-cut.grass
'He cut down the surrounding weeds.'

1.1.8 Many subjects, few objects (ri-). Although the pronominal prefixes indicate number, there is an additional prefix, ri-, which indicates that the number of agents in

the clause is considerably greater than the number of patients (19, 20).

(19) o kaso ma gudai ina-<ri>-goliokau
ART dog POS many they.them-<FO>-bit
'The pack of dogs bit the (few) people.'

(20) ma koana guuna woboana de wi-<ri>-wako
ART king that he.came and they.him-<FO>-reject
'The king came and they rejected him.'

1.1.9 Continuative (ta-). This prefix occurs only in the reduplicated form in Tabaru. The ta- prefix appears to have come from a form tV- which is no longer productive (Fortgens 1928:385).

(21) mia-ta-<ta>-ngodomokasi una woboa
we.it-RDP-<PRO>-meal.CON he he.came
'While we were still eating, he came.'

(22) cedeke ka mo-ta-<ta>-iete de
EXC but she-RDP-<PRO>-laugh EXC
'Hey!, she just keeps laughing.'

1.1.10 Habitual (taba-). To indicate that someone is in the habit of performing a particular action, the prefix taba- is used (23, 24). This prefix is rarely used in Fortgens' folk stories (1928), a search turning up only one example, although there are examples in his grammatical notes.

(23) ni riaka ni-<taba>-sosongene
your older.sisters they.you-<HAB>-RDP.kill
'Your older sisters repeatedly kill you.'

(24) ami kule mi-<taba>dongosama
her evil.spirit her-<HAB>-APl.possessed
'She is continually possessed by an evil spirit.'

1.1.11 Applicative 3 (g-). An example of the prefix g- on vowel initial verb root is:

ari 'cry'
gari 'cry for'

In the examples below (25, 26), the g- prefix occurs before a voiceless consonant:

parene 'rise'
taulu 'cook'

In these cases, the first consonant of the root becomes voiced and the g- is dropped. For further explanation of the voicing of initial consonants and the deletion of the g-, see Chapter 4 on phonology.

(25) ...la uwa o mokumoku yo-<g>-barene
so NEG ART waves they-<AP3>-rise
'...so that the waves don't rise upon us.'

(26) dua asa naino noi-<g>-daulu
moment then you.it.here you.p-<AP3>cook
'Then you come here and cook for me.'

1.2 Suffixes

The suffixes in Tabaru are indicators of tense, aspect, polarity and spatial movement. The following is a template showing suffixation on the Tabaru verbs:

VERBAL SUFFIX TEMPLATE

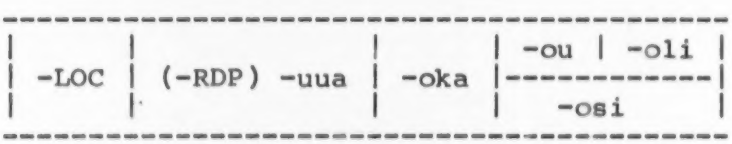


FIGURE 12

The chart (above) shows not only the relative order of occurrence of the suffixes, but also the co-occurrence restrictions. For example, of all the suffixes, only the negative suffix -uwa can be reduplicated. This is indicated by the occurrence of the -RDP morpheme inside the same box as -uwa. The suffix -osi cannot co-occur with -ou or -oli and for that reason the former suffix is located below the latter suffixes in the diagram; the non-linear order indicates that they cannot co-occur.

1.2.1 Locational suffixes (-LOC). The locational suffixes indicate location (27) or direction of movement (28). Figure 13 below is a list of these suffixes.

(27) wogoger-<ie> o mia ma dateoko
he.sit-<LOC> ART monkey POS side
'He sat above, beside the monkey.'

(28) tooluku tonisakar-<uku>
I.don't.want I.you.push-<LOC>
'I don't want to push you down.'

Spatial deixis is very complex in Tabaru. Yoshida (1980) says that in Galela (a language closely related to Tabaru) various location and direction-of-movement words and affixes can be combined to produce over one hundred different indications of location, movement and destination. Tabaru (as well as the other North Halmaheran languages) has a similar spatial deixis system. As complex as this system may be, when used as a suffix on the non-deictic verbs in Tabaru, only one location suffix occurs on any given verb. The following is a list of the Tabaru directional suffixes and their meanings.

LOCATIVE VERBAL SUFFIXES

-ie	'up'	-uku	'down'
-isa	'landward'	-oko	'seaward'
-ino	'toward speaker'	-ika	'from speaker'

FIGURE 13

Tabaru also uses these forms as independent verbs to indicate direction of movement (29 & 30).

(29) ngona na-<ino>
you you-<this.way>
'You come here!'

(30) o giajooungu o bonganoka ya-<isa>
ART couple ART forest.LOC they-<landward>
'A couple went (landwards) to the forest.'

1.2.2 Negation (-uwa). The suffix -uwa indicates the negation of the verb to which it attaches.

(31) alele ka noiakun-<uwa>
thanks but you.me.eat-<NEG>
'Thank you, you didn't eat me!'

(32) ngaro imakuata ma yaakun-<uwa>
although it.REF.strong but it.it.could-<NEG>
'Although he (crocodile) strengthened himself,
he couldn't do it.'

1.2.3 Reduplication, emphatic negation (-RDP). The negative suffix -uwa is the only suffix which is reduplicated. It indicates emphatic negation (33, 34).

(33) ngoi neenaka tomagogere ma taakun-uwa-<uwa>-ou
I this I.REF.live but I.it.take-NEG-<RDP>-PER
'I live here but I just can't take it anymore!'

(34) taakukun-uwa-<uwa>-ou tooara
i.it.able-NEG-<RDP>-PER I.run
'There's no way I can run.'

1.2.4 Past tense (-oka). Tabaru has a two-tense system: past tense, which is marked by the suffix -oka (35, 36), and a non-past which is unmarked. It is assumed at this time that the second verb in (36) takes on the past tense from the first verb.

(35) yapidil-<oka> ka dai o Tarinateka
they.it.release-<PST> only east ART Ternate.LOC
'They let it go east of Ternate.'

(36) wapait-<oka> womimake
he.it.dig-<PST> he.her.meet
'He dug there and he found her.'

1.2.5 Already (-ou). The -ou suffix in Tabaru conveys the meaning of 'already' (37, 38).

(37) yoli-<ou> o woaka
they.return-<ALR> ART house.LOC
'They already return home.'

(38) o gaale ka wododorokut-<ou>
ART snail only he.RDP.fill-<ALR>
'He is already putting snails in.'

1.2.6 Repetitive (-oli). The affixation of -oli on a verb indicates repeated action (39, 40).

(39) tomasouru de taisa toli-<oli>
I.REF.medicine and I.it.landwards I.return.-<REP>
'I got medical help and then I returned again.'

(40) asade yotagi-<oli>
then they.go-<REP>
'Then they left again.'

1.2.7 Continuative (-osi). When -osi is affixed to verbs, they have a continuative meaning (41, 42).

(41) ai ngowaka geena wosikol-<osi>
my child that he.schools.<CON>
'My son is still in school.'

(42) goona o naoko yoodom-<osi>
they ART fish they.eat-<CON>
'They are still eating fish.'

2 Nominal morphology

Nominal morphology in Tabaru is relatively simple. There is no singular-plural distinction. Tabaru nouns can be both prefixed and suffixed, but unlike the verb affixation, noun affixation is not obligatory. The following is a template of noun affixation:

NOUN AFFIXATION TEMPLATE

RDP-	NOUN	-LOC

FIGURE 14

2.1 Reduplication

Reduplication on Tabaru nouns can be divided into two categories: one which occurs on underived nouns and one which occurs on derived nouns.

On the underived nouns, the RDP- prefix has two possible meanings: that of a collective as in examples (43) and (44) or that of resemblance to the original noun as in examples (45) and (46). When RDP- prefix occurs on nouns derived from verbs, it takes on the meaning of instrument as in examples (47 and 48). The phonological process of the nominalizing (NML) prefix g- is the same as that of the homophonous -applicative (AP3). (See the phonology section under the consonant voicing rule and g- deletion.)

(43) o saeke — o sae-saeke
ART head, leader — ART (all) leaders

(44) o ngowaka — o ngowa-ngowaka
ART children — ART descendents

(45) o ngootiri — o ngo-ngootiri
ART boat — ART miniature boat

(46) o nyawa — o nyawa-nyawa
ART man — ART doll

(47) una wotike o <ba>-baasana
he he.look.for ART <RDP>-NML.beat
'He looked for a cudgel.'

(48) awi aako ma <da>-dawi ma itiwa ma
his eyes POS <RDP>-NML.wink only they.close only
'His eyelashes felt like closing.'

2.2 Locational suffixes

Tabaru nouns take the same location suffixes (figure 13) as the verbs. Examples (49) and (50) demonstrate the usage of two of these suffixes. Besides these locational suffixes, nouns can take an additional suffix -oka, which does not occur on verbs. (There is an -oka suffix on verbs, which as we have seen indicates past tense.) This suffix is exemplified in (51) and (52), and has the meaning 'in' or 'from'. (See the section on phonology for a further description of suffixation and stress). For example:

(49) o aker-<ika>
ART river-<LOC>
'to the river (from speaker)'

(50) o ngootir-<uku> woboano
ART boat-<LOC> he.come.LOC
'He came into the boat.'

(51) ngomi o Tabaru-<oka> jou
we ART Tabaru-<LOC> sir
'We are from Tabaru, sir.'

(52) womasioru o ngootiri o ngoot-<oka>
he.REF.paddle ART boat ART ocean-<LOC>
'He paddles the boat in the ocean.'

3 Derivational morphology

In the discussion of derivational morphology, we will examine here only that morphology which changes word classes. Other derivational affixes such as the applicative have been discussed above under verbal morphology.

3.1 Verb derivations

Verbs are derived from nouns by the addition of a Ø morpheme in the following manner:

$$[[\emptyset]_V\ [\ ]_N]_V$$

The pairs of examples below are noun followed by verb:

(53) waasa o mokuru geena
he.it.bring ART betel.nut that
'He brings the betel nut.'

(54) wonitimaiu la wo-<mokuru>
you.him.give so he-<betel-nut>
'You give some to him so he can chew betel-nut.'

(55) awi singina yaowa
his heart it.good
'His heart is good.'

(56) woki-<singina> ma esa de ma dea
he.them-<heart> POS mother and POS father
'He considered his mother and father.'

3.2 Noun derivations

Nouns are derived from verbs by the addition of the g- prefix in the following way (55) and (56):

$$[[g\text{-}]_N\ [\ ]_V]_N$$

As was seen above in the discussion of the applicative (AP3) prefix on the verb (1.1.11 above), the nominalizing prefix (NML) g- on the verb appears overtly only on vowel-initial verb roots. On verbs with consonant-initial roots, the g- causes voicing in the voiceless consonant and the g- is not connected to a syllable and is therefore dropped. This is exemplified below.

g- + temo --> gtemo --> gdemo --> demo

For a further discussion of this issue, see Chapter 4.

(57) una wosidemo to ena ma <g>-demo
he he.AP2.AP3.say POS it POS <NML>.say
'He spoke in his own words.'

(58) ma <g>-akuru o mongoo butanga bolo
POS <NML>-is.tall ART fathom six only
'Its height is only six fathoms.'

NOTES

[1]Within the framework of Relational Grammar, Perlmutter (1978) discusses the notion of initial unergativity which he ties to the verbal concepts of activity and involuntary processes. The notion of unaccusative is tied to adjectival verbs or those with patient as an initial nuclear term. An explanation of the syntax of this phenomenon is beyond the scope of this paper. However, it should be noted that the semantic distribution of Tabaru verbs does not adhere to Perlmutter's proposed pattern. In Tabaru, temporary state or involuntary action are unaccusative predicates as expected, but attributive predicates (as well as active intransitives, as in examples 1 and 2) are unergative.

[2]It appears that for the sake of completeness, Fortgens (1928) included within his grammatical notes forms that rarely occur, as there are few if any examples of these within the folk tales he collected, and none in my field notes.

CHAPTER 4
PHONOLOGY

The Tabaru language has been described by Fortgens (1928). However there is much left to be said about stress, syllable structure, and phonological rules.

We will describe the phonemic inventory, stress placement, syllable structure, lexical and postlexical rules, and propose a number of analyses. The purpose of this section is to describe Tabaru phonology and determine the description which proves to capture important generalizations about Tabaru, within a framework that is both predictive and restrictive in its generalizations about human language.

1 Inventory of segments

We will be showing here the inventory of the underlying segments in Tabaru by the examination of contrasts, and discuss the existence of phonetic semivowels. We use here the conventions of the Indonesian orthography. Except for the following symbols shown below, the symbols used have their regular IPA values.

Indonesian		International Phonetic Alphabet
ny	=	ɲ
ng	=	ŋ
j	=	dʒ
r	=	ɾ
c	=	tʃ

The segments f, h, ny, c, and j are rare in Tabaru and occur only in borrowed words. In the following chart they are enclosed in parentheses.

CONSONANT CHART

	LABIAL	ALVEOLAR	ALVEOPALATAL	VELAR	GLOTTAL
VOICELESS STOP	p	t	(c)	k	
VOICED STOP	b	d	(j)	g	
FRICATIVE	(f)	s			(h)
NASAL	m	n	(ny)	ng	
LATERAL		l			
VIBRANT		r			

VOWEL CHART

	FRONT	CENTRAL	BACK
HIGH	i		u
MID	e		o
LOW		a	

1.1 Consonants

Stops in Tabaru are bilabial, alveolar and velar. There are both voiced and voiceless stops at each point of articulation.

Contrasts:

[p] : [b]	'poa	'pour out'
	'boa	'come'
	ta'padu	'oppress'
	ta'bau	'I borrow it'
[t] : [d]	'towo	'give food'
	'dowo	'smoke'
	'uekata	'wife'
	kaka'daku	'above'
[k] : [g]	'kakuru	'tall'
	'gakuru	'corn'
	'otaka	'accompany'
	'tagi	'go'

The fricatives are labial, alveolar and glottal, all of which are voiceless.

[f] : [s]	'fara	'type'
	'saara	'snake hole'
	'ngafi	'a fish'
	'asiri	'swallow'
[f] : [h]	'faja	'dirty'
	'haga	'hold, embrace'
	'lefo	'writing'
	'lahi	'request'
[s] : [h]	'sago	'branch'
	'faga	'hold, hug'
	'asiri	'swallow'
	'lahi	'request'

The only affricates in Tabaru are alveopalatal voiced and voiceless, which occur largely in borrowed words.

[d] : [j]	'dou	'foot'
	'jou	'mister'
	'bido	'betel nut'
	'ija	'buy'
[t] : [c]	ti'osaka	'wild'
	cio'tiono	'very small'
	'toto	'seed'
	co'coro	'funnel'
[c] : [j]	'cobo	'grab'
	'jobo	'depart'
	bi'cara	'say'
	'ija	'buy'

There are four voiced nasals: bilabial, alveolar, alveopalatal and velar.

[m] : [n]	a'mia	'ours'
	a'nia	'yours'
	'mao	'feel'
	'nao	'lining'
[n] : [ny]	'nauru	'husband'
	'nyawa	'person'
	ma'nai	'friend'
	ma'nyanyi	'incense'
[ny] : [ng]	'nyiara	'nut tree'
	'ngia	'snake'
	ma'nyanyi	'incense'
	'manga	'our (incl)'

There are two voiced liquids: an alveolar lateral and an alveolar flap.

[l] : [r]	'lio	'return'
	'rio	'footprint'
	'ali	'chest'
	'ari	'weep'

1.2 Vowels

Tabaru has five vowels: a, e, i, o, and u. Fortgens (1928:301-304) describes a variation in the [-high] vowels dependent upon environment. Other researchers in the North Halmaheran languages used diacritics on the vowels in their descriptions (see van Baarda 1895 and Hueting 1908) to show similar variations. It is unclear what his description actually refers to. It appears that Fortgens wished to be as careful as his colleagues in describing the sound distinctions, yet he was convinced that the significance of these differences was not great enough to include them in an orthography. When Dr. Adriani (van der Veen 1915:11) visited Halmahera to review the situation for himself, he did not hear the distinctions. As can be seen below in the contrasts, there are only five vowels in the Tabaru language.

Contrasts:

[a] : [e]	'adana	'braid'
	'edete	'grandparents'
	'baili	'remainder'
	'beika	'try
[a] : [o]	'alu	'breastbone'
	'olu	'flood'
	'sauku	'warm'
	'souru	'medicine'
[e] : [i]	'edete	'braid'
	'iete	'laugh'
	'dea	'father'
	'dia	'knife'
[o] : [u]	'ouku	'bamboo container'
	'uunu	'read'
	'alo	'cold'
	'alu	'breastbone'
[i] : [u]	'iete	'laugh'
	'uere	'plait'
	'wisa	'bark girdle'
	'wusi	'comb'

1.3 Semivowels

The status of the semivowel is unclear in Fortgens' analysis (1928:305-306, 315). A summary of his views follows:

1. There is only one semivowel, the y (1928:304).
2. The y is rarely used (1928:305).
3. The words in which the y occurs are borrowed (1928:305).
4. Y occurs in pronominal verb prefixes (1928:305,337).
5. An epenthetic w appears between a labial vowel and another vowel (1928:305).
6. Where the Tabaru language has a w, the Ternate language has an f (1928:315).

It is unclear why Fortgens would make a statement like (1) above when he goes on to discuss the w in such statements as (6). Perhaps his reasoning is that because, in the Tabaru-adu dialect, the w is deleted (1928:305), it need not be considered. He correctly states that y is rarely used in Tabaru (2) and that it occurs in borrowed words (3); however he then states that y occurs in the verb pronominal prefixes (4). It is improbable that a borrowed segment should appear in borrowed words and also in the pronominal verb prefixes. His clearest remark about semivowels is (5) regarding an epenthetic w.

Up to this point we have been using the y and w in the examples in this paper. It should be noted that [y] and [w] are underlyingly /i/ and /u/ and that the transcription as y and w has been used for the sake of clarity. The symbols y and w will not be used in the formal derivations because the semivowels can be determined by general convention, as seen below.

It is clear that the phonetic segments [y] and [w] appear in the Tabaru language. With the introduction of the notion that the syllable is a hierarchic structure, we are barred from using the feature [syllabic]. This feature was used to distinguish the semivowels from the [+high] vowels. The phonetic realization of the [-consonantal, +high] segments is determined by syllable structure, which contrasts to Fortgens' unclear analysis of these segments.

2 Syllable structure

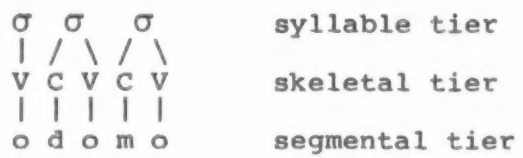
The notion of syllable has had a rather difficult history. Although embraced by structuralist theory (Bloomfield 1933, Gleason 1961, Pike 1967), it was rejected in early generative theory (Chomsky and Halle 1968). More

recently, research (Kahn 1976, Clements and Keyser 1983, Harris 1983) has indicated the importance of the notion of the syllable. As seen above, an adequate theory of stress will use the notion of the syllable.

We assume here the notion of the syllable, and that CV is the core syllable in human language. All other syllable types can be derived from the core syllable by the application of one or both of the following rules[1] (Clements and Keyser 1983:28, 29):

a. delete syllable initial C
b. insert syllable final C

The Tabaru language makes use of only rule (a) with the result that there are only CV and V syllables. We will assume with Clements and Keyser (1983) a three-tiered model of the syllable. The following examples illustrate this model.



Harris (1983:76) discusses syllable structure in terms of rules:

> One characteristic of these rules that clearly sets them apart from garden-variety phonological rules is that syllable structure rules do not change features. Rather they belong to the class of rules that assign features and/or prosodic structure that is lexically unspecified.

Languages tend to have syllables of the CV type. If a language has only one syllable type, it will be CV. The onset rule states that languages will maximize onsets in their syllables, that is, they will produce onsets to syllables wherever possible. The Tabaru language is no exception to this. A link between a segment and a C-element produces a consonant, and a link between a segment and a V-element produces a vowel. Thus if a [+high, -consonantal] segment is linked to a C-element, a consonant ([y] or [w]) is produced, but if it is linked to a V-element, the result is a vowel, ([i] or [u]). Clements and Keyser (1983:32)

examine the notion of linking the segmental tier with elements on the skeletal tier:

> Unless otherwise stipulated in the grammar or lexicon of a given language, V-elements of syllable structure are freely allowed to dominate [-consonantal] segments, and C-elements are freely allowed to dominate both [+consonantal] segments and [+high, -consonantal] segments. Other associations are possible only when admitted by language specific rules.

We assume that "freely allowed to dominate" means it produces a well-formed syllable. In other words, [+high, -consonantal] segments can dominate both C-elements and V-elements in the syllable structure. If it is linked to a V-element, it is realized phonetically as a vowel, but if it is linked to a C-element, it is realized phonetically as a semivowel.

It is then necessary to determine the criterion used to link the [+high, -consonantal] segments to the proper element on the skeletal tier. Given that languages maximize onsets, Clements and Keyser (1983:37) state the following:

> a. Syllable-initial consonants are maximized to the extent consistent with the syllable structure conditions of the language in question.
>
> b. Subsequently, syllable-final consonants are maximized to the extent consistent with the syllable structure conditions of the language in question.

Based upon this notion, we will assume that a language will attempt to link any [+high, -consonantal] segment to a C-element if at all possible, but if it is unable to link to a C-element, because of language-specific restrictions, it will link to a V-element. Because there are no syllable final consonants in Tabaru, we can ignore (b) above.

Let us examine the process of maximizing the onsets, using the Tabaru words [ngununu], 'nose' and [wekata], 'wife' as examples.

1. Link C-elements to [+consonantal] segments.

```
   C   C   C                C   C
   |   |   |                |   |
  ng u n u n u              u e k a t a
```

2. Link V-elements to [-consonantal, -high] segments.

```
   C   C   C                V C V C V
   |   |   |                | | | | |
  ng u n u n u              u e k a t a
```

3. Wherever possible, link [-consonantal, +high] segments to C-elements.

```
   C   C   C                C V C V C V
   |   |   |                | | | | | |
  ng u n u n u              u e k a t a
```

4. The remainder of the [-consonantal, +high] segments are linked to V-elements.

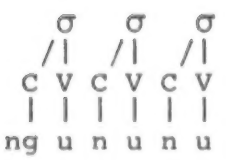
```
   C V C V C V              C V C V C V
   | | | | | |              | | | | | |
  ng u n u n u              u e k a t a
```

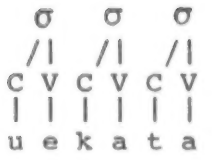
It should be noted here that the above rule accounts for the phonetic [u] in [ngununu] and the phonetic [w] in [wekata]. The [+high, -consonantal] segments which are linked to C-elements are the semivowels, and those linked to V-elements are vowels.

The linking of the skeletal tier and the segmental tier is formed by the following steps (Clements and Keyser 1983:38):

a. V-elements are prelinked to σ's
b. C-elements to the left are adjoined one by one as long as the configuration resulting at each step satisfies all relevant syllable structure conditions.
c. Subsequently, C-elements to the right are adjoined in the manner described in (b) above.

Using the words from the examples above in linking the skeletal to the syllable tier we have the following:





Part (c) of the above algorithm does not apply because there are no closed syllables in the Tabaru language. The syllabification rules apply to the underlying form (before affixation) and after each new morpheme is added.

Tabaru has a simple syllable structure in which it is not necessary to write rules for the occurrence of segments in onset or nucleus positions. These fall out naturally from theoretical considerations: the theory predicts only [+high, -consonantal] segments and [+consonantal] segments can occur in an onset, and only [-consonantal] segments can occur in a nucleus. The syllable types in Tabaru are all either V or CV. The V syllable consists of a segment linked to a V-element, which is prelinked to a syllable on the syllable tier. This is seen in the first syllable of the following word.

'his'

In Tabaru, an onset is a segment which is linked to a C-element on the skeletal tier. The onset is sometimes complex as in the case of the borrowed segments ny (ɲ), c, (tʃ) and j (dʒ). See examples 1-5.

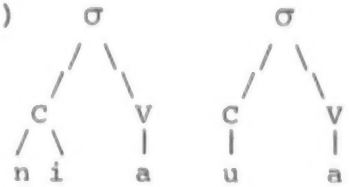
(1)

'white cockatoo'

(2)

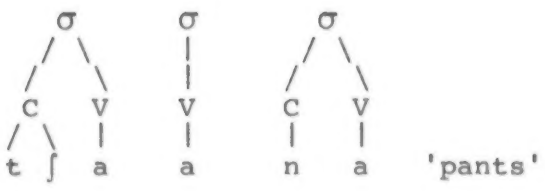
'my'

(3)



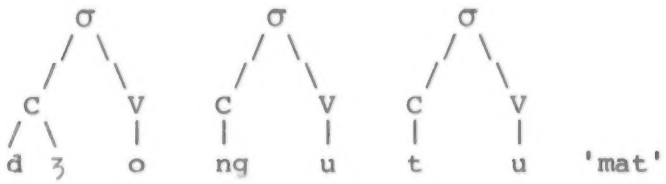
'person'

(4)



'pants'

(5)



'mat'

3 Affixation

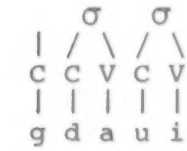
3.1 Rules associated with the g- prefix

The g- prefix is affixed to verbs as either an applicative or derivational morpheme as seen in Chapter 3. When it is added to vowel-initial roots, it produces derived forms like those below:[2]

ari	'cry'	gari	'cry for'
asa	'bring'	gasa	'bring with'

When the g- is added to words with initial voiceless stops, the derivation is more complex. There is a voicing rule which causes the voiceless stop to become voiced. The skeletal and syllable tiers are then added and any unattached segments are not realized phonetically, as seen in the example below:

taui	-->	gtaui	g- prefixation
gtaui	-->	gdaui	voicing rule



syllable structure

dawi — surface form

There is some question as to whether stems with initial voiced and fricative consonants underlyingly are affixed with the g- prefix. If such is the case, then the g- never appears on the surface. The following is an example of how

nouns are derived from verbs by addition of the g- prefix, and how other nouns are further derived from these new nouns by reduplication.

pikiri	'to strain'	
bikiri	'sieve'	nominalizing rule
bibikiri	'strainer'	reduplication

If we assume, by analogy, that verbs with voiced initial consonants underlyingly also undergo the same nominalizing rule (without overt effect) before they reduplicate, then g- is affixed to them in the derivation of the related words.

3.2 Reduplication

The reduplication morpheme RDP- occurs both on verb prefixes and suffixes, as was seen in the section on morphology. Because reduplication functions in a similar manner on both prefixes and suffixes, we will discuss them together here. Reduplication is accomplished by the reproduction of the syllable. McCarthy and Prince (1986) argue against using a skeletal tier template, and in favor of using a syllable tier template. We follow them here in assuming reduplication occurs first of all on the segmental tier (melody), and then the syllable tier is reproduced. Finally, the syllable tier is mapped onto the segmental tier to the fullest possible extent within the rules of the language. For example:

	Reproduce Segments	Reproduce Syllable	Map Syllables to Segments
σ σ	σ σ	σ σ σ σ	σ σ σ σ
/|/|	/|/|	/|/|	/|/|/|/|
tagi -->	tagitagi -->	tagitagi -->	tagitagi

The segmentals are mapped (linked) from left to right for prefixes and from left to right for suffixes.

3.3 Vowel deletion

There is a process in Tabaru that deletes suffix vowels when suffixes are added to roots or to existing suffixes; for example:

tagi + uua --> tagiua
lio + osi --> liosi

It appears that the following two rules are needed to account for deletion:

$$V \rightarrow \emptyset \; / \; \underset{[+high]}{V} \; _____ \; \underset{[+high]}{V}$$

$$V \rightarrow \emptyset \; / \; V \; _____ \; \underset{[-high]}{V}$$

Clearly these two rules are in fact the same process. They are both deletion rules and they both affect suffixation. Only the height of the vowel is different between the rules. In David Tuggy's article (1979), he discusses various interpretations of angle brackets and their uses. He defines an interpretation of angle brackets which works both in terms of traditional interpretation and also in cases where angle brackets could not apply by traditional interpretation. This would allow for the collapse of the above two rules. Tuggy (1979:114) says:

> The notion is intuitively quite simple: The if-bracket states a condition and the then-bracket expresses a requirement that is dependent on that condition. If the condition in the if-bracket is fulfilled, then so must be the requirement in the then-bracket. However, if the condition in the if-bracket is not fulfilled, then it does not matter whether the requirement in the then-bracket is satisfied or not. Thus, on the first pass you check to see if the condition in the if-bracket and its corresponding requirement in the then-bracket are satisfied. If they are, you may apply the rule. Then on the second pass you check to make sure that the if-condition is not satisfied. If it is not, you may apply the rule, whether or not the then-requirement is satisfied.

Following Tuggy we will mark the if-bracket $\langle\rangle_i$. Using this notation, the above rules collapse as follows:

$$\underset{[\langle +high \rangle]}{V} \rightarrow \emptyset \; / \; V \; _____ \; \underset{[\langle +high \rangle_i]}{V}$$

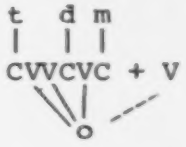
The above rule reads: If an environment vowel X is [+high] and is followed by vowel Z which is also [+high], then vowel

Y is deleted; if environment vowel X is [-high] and is followed by any vowel Z, then vowel Z is deleted.

3.4 Vowel epenthesis

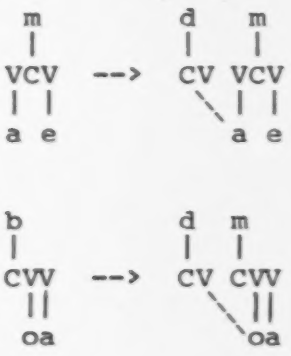
In the discussion of the metrical structure above, it was noted that there are Tabaru roots which in underlying form end in consonants. Tabaru however, does not allow closed syllables in the surface form, and rather than delete the final [+consonantal] segment, a vowel is epenthesized on the end of the word. This works by the addition of a V to the skeletal tier of the root which is then linked to the previous V on the segmental level. For example:

toodom 'I eat'



3.5 The dV- prefix

The linking of the unspecified vowel in the dV- prefix, as a matter of general convention, links the V to the nearest segment. This linking is illustrated in the following example:



In contrast, the examples in the previous section link to the right.

3.6 Stress

The rules for stress, as discussed in Chapter 2, are:

1. Construct right-branching, left-headed feet.
2. Construct right-headed words.

Derivations including stress are presented at the end of this chapter.

3.7 Pronominal prefixes

There are irregularities within the pronominal transitive verb prefixes in the Tabaru language which are beyond the scope of this study, although they do warrant further investigation. It is clear that these forms are derived through word formation rules and phonological processes[3]. It is not clear beyond doubt that these are prefixes. The verbs never occur without these, even within a relative clause. This is evidence for their status as prefixes. Argument against prefix status is the fact that there are no phonological processes between them and the verb as there are for clear cases of affixation. For the following derivations, we will assume that the underlying forms of the transitive pronominal verb prefixes deviate from the surface forms only in the formation of semivowels.

4 Lexical phonology

Lexical phonology as conceived by Mohanan (1986) and also by Kiparsky (1982) has its appeal in the ability to solve problems relating to boundaries and in positing a level of representation which is similar to the phonemic level of structuralist theory.

Within lexical phonology, the lexicon contains both words and affixes. Phonological rules may apply in two domains: the lexical module and the postlexical module. The rules are ordered and specified as to which module(s) they apply in. There may also be a number of strata within the lexical module. The phonological rules must be specified as to which strata the rules apply. The following figure exemplifies this model:

A MODEL OF LEXICAL PHONOLOGY

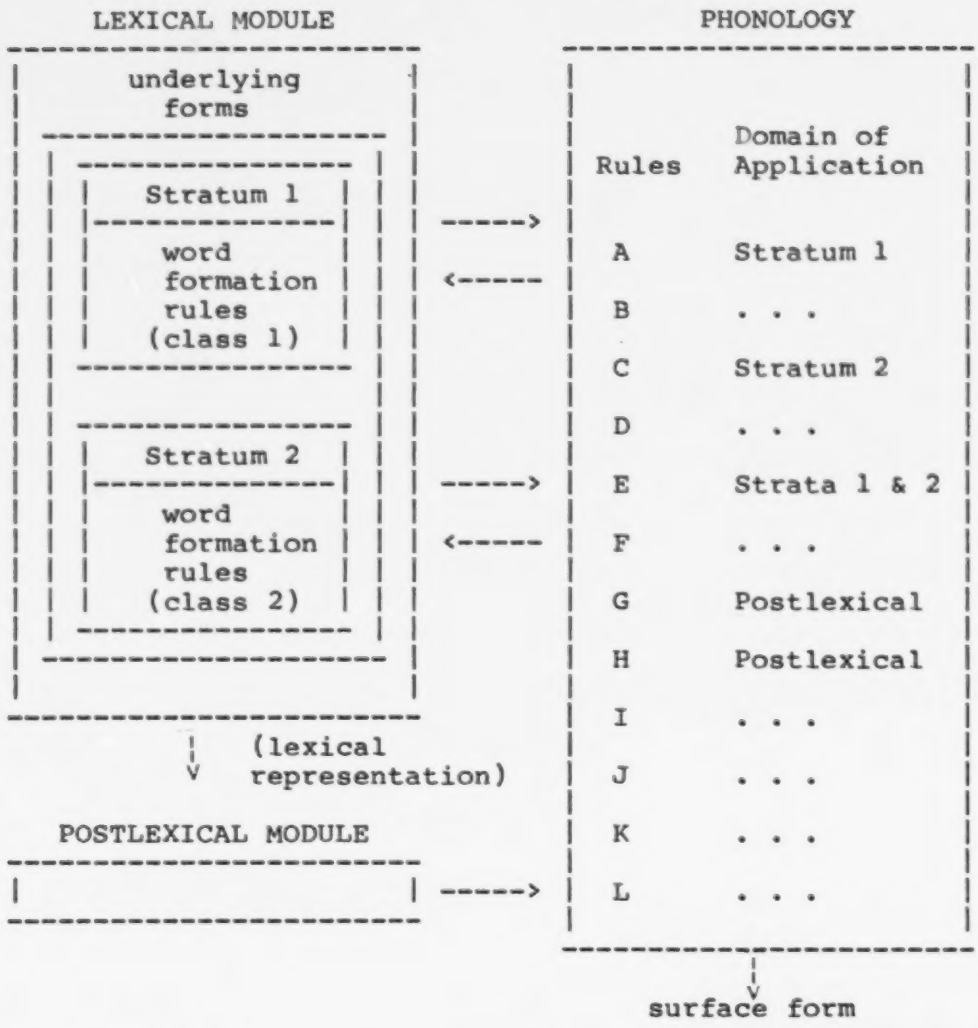


FIGURE 15

The word formation rules apply affixes to words. Each word formation rule is located within a particular stratum. After the words are affixed, they move to the phonological module where the appropriate rules for that stratum apply. The words then return to the stratum from which they came. If there is no further affixation on that stratum, the words then move on to the next stratum or module. Whenever a word moves from one stratum or module to another, the brackets indicating morphological information are erased, as exemplified by the Tabaru derivations below.

5 Lexical phonology of Tabaru

In the lexical module, Tabaru has two strata. Within the first stratum, suffixation occurs. The second stratum is reserved for prefixation. The necessity of two strata is clear from the fact that vowel deletion does not apply after prefixation, as it does after suffixation, even though the structural description may be met. (See Sect. 3.3 above on vowel deletion.)

The list below enumerates the phonological rules in Tabaru and indicates in which strata the rules apply. As has been noted above, syllabification applies wherever possible. This is also true of linking. In the derivations, these will be shown only where they crucially apply.

Tabaru Phonological Rules:

1	syllabification	general convention
2	voicing	C --> C / C --- [+voice] [+voice]
3	linking	linking
4	vowel deletion	V --> Ø / V --- [<+high>] [<+high>i]
5	stress assignment	1. build right-branching, left-headed feet. 2. build right-headed words
6	V epenthesis	Ø --> V / C ---]
7	C epenthesis	Ø --> C / V --- V

Phonological Rules:		Domain of Application:
1	syllabification	level 1, 2, postlexical
2	voicing	level 2
3	linking	level 1, 2, postlexical
4	vowel deletion	level 1
5	stress assignment	postlexical
6	V epenthesis	postlexical
7	C epenthesis	postlexical

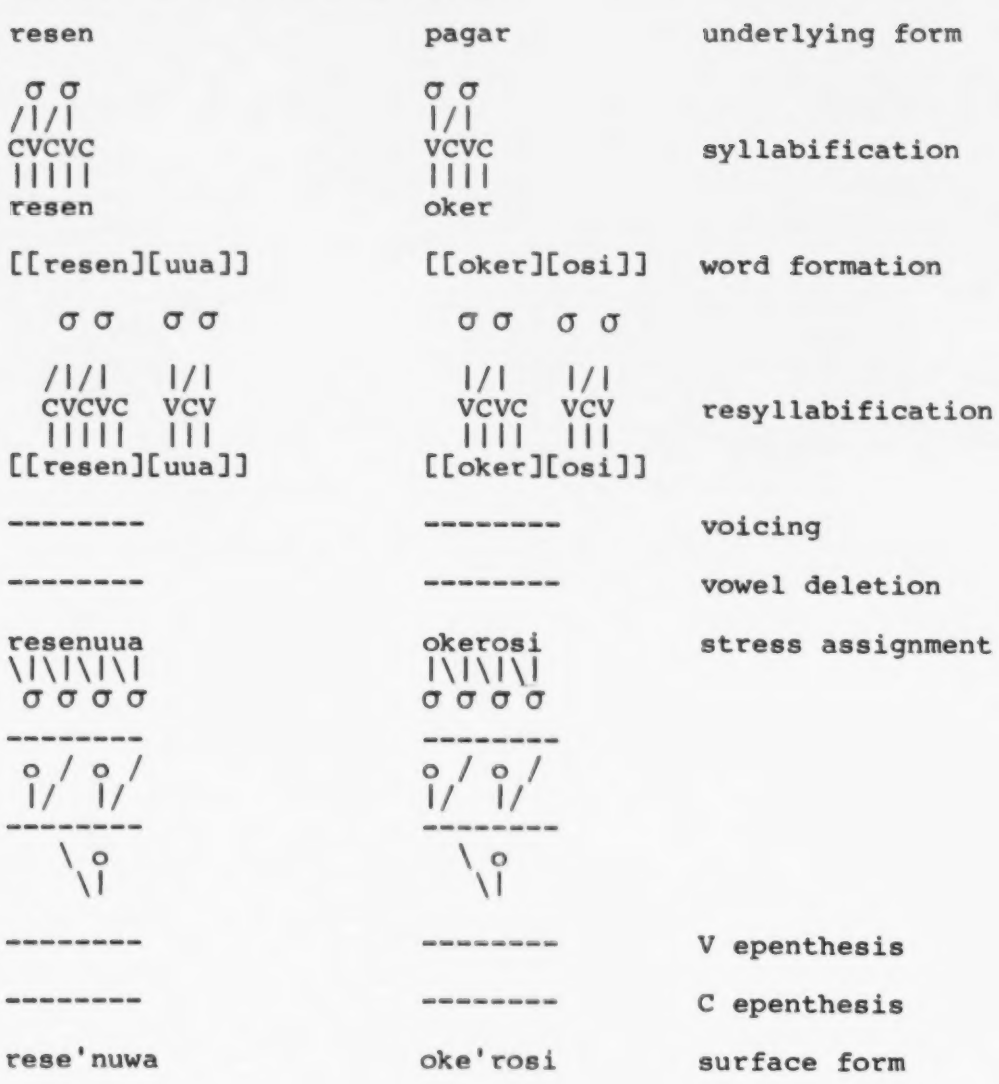
6 Derivations

Let us examine some sample derivations which are representative of how rules apply to Tabaru data.

Consonant final stem without suffixation:

oker	pagar	underlying form
σ σ |/| VCVC |||| oker	σ σ /|/| CVCVC ||||| pagar	syllabification
----	-----	voicing
----	-----	vowel deletion
oker |\\\| σ σ ---- o / |/ ---- o | |	pagar \\\|\\\| σ σ ----- o / |/ ----- o | |	stress assignment
k r | | VCVCV | | o e	p g r | | | CVCVCV \\\| a	V epenthesis
k r | | VCVCV | |/ o e	p g r | | | CVCVCV \\\|/ a	linking
σ σ σ |/|/| VCVCV ||||| okere	σ σ σ /|/|/| CVCVCV |||||| pagara	resyllabification
-----	------	C epenthesis
'okere	'pagara	surface form

Affixed consonant final stem.



Vowel final stem without affixation.

tagi	sofala	underlying form
σ σ /|/| CVCV |||| tagi	σ σ σ /|/|/| CVCVCV |||||| sofala	syllabification
----	------	voicing
----	------	vowel deletion
tagi |||| σ σ ---- o / |/ ---- o |	sofala ||||| σ σ σ ------ o o / | |/ ------ \ o |	stress assignment
----	------	V epenthesis
----	------	C epenthesis
'tagi	so'fala	surface form

Vowel final stems affixed with [+high] initial vowel suffixes.

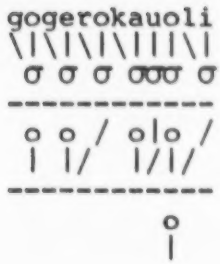
tagi	sofala	underlying form
σ σ /|/| CVCV |||| tagi	σ σ σ /|/|/| CVCVCV |||||| sofala	syllabification
[[tagi][uuа]]	[[sofala][uuа]]	word formation
----------	----------	voicing
[[tagi][ua]]	[[sofala][ua]]	vowel deletion
σ σ σ /|/| CVCV |||| [[tagi][ua]]	σ σ σ σ /|/|/| CVCVCV |||||| [[sofala][ua]]	resyllabification
tagiua ||||| σ σ σ	sofalaua |||||| σ σ σ σ	stress assignment
------ o o / | |/ ------ \ o ||	-------- o / o / |/ |/ -------- \ o ||	V epenthesis
------	--------	C epenthesis
ta'giwa	sofa'lawa	surface form

Vowel final stem affixed with [-high] initial vowel suffixes.

tagi	kula	underlying form
σ σ /|/| CVCV |||| tagi	σ σ /|/| CVCV |||| kula	syllabification
[[tagi][ou]]	[[kula][osi]	word formation
----------	----------	voicing
----------	[[kula][si]]	vowel deletion
σ σ σσ /|/| || CVCV VV |||| || [[tagi][ou]]	σ σ σ /|/| /| CVCV CV |||| || [[kula][si]]	syllabification
tagiou |\\||| σ σσσ ------ o /o / |/ |/ ------ \ o \\| ------	kulasi |\\||| σ σ σ ------ o o / | |/ ------ \ o \\| ------	stress assignment V epenthesis
------	------	C epenthesis
tagi'ou	ku'lasi	surface form

Multiple suffixation is exemplified in the following derivation. This derivation is important because it crucially depends upon the resyllabification convention for proper output. The resyllabification after the addition of the suffix -oli would link the [u] with a C-element if it were not for the prohibition that links with V-elements are not erased when forms are resyllabified. In other words, the output of the derivation is a vowel, not a semivowel as we would expect if resyllabification erased all links between the segmental and the syllable tiers.

goger	underlying form
σ σ /|/| CVCVC ||||| goger	syllabification
[[goger][oka]]	word formation
----------	voicing
----------	vowel deletion
[[goger][oka][ou]]	word formation
----------	voicing
[[goger][oka][u]]	vowel deletion
σ σ σ σ σ /|/| |/| | CVCVC VCV V ||||| ||| | [[goger][oka][u]	resyllabification
[[goger][oka][u][oli]]	word formation
----------	voicing
----------	vowel deletion
σ σ σ σ σ σ σ /|/| |/| | |/| CVCVC VCV V VCV ||||| ||| | ||| [[goger][oka][u][oli]]	resyllabification



stress assignment

V epenthesis

C epenthesis

gogerokau'oli

surface form

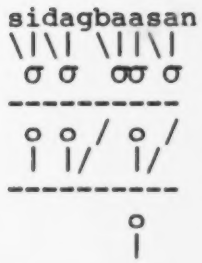
Examples of the dV- and g- prefixation.

omas	ari	underlying form
σ σ |/| VCVC |||| omas	σ σ |/| VCV ||| ari	syllabification
[[dV][omas]]	[[g][ari]]	word formation
----------	----------	voicing
σ σ σ /| |/| CV VCVC || |||| [[dV][omas]]	σ σ |/| C VCV | ||| [[g][ari]]	resyllabification
d m s | | | CVVCVC | | o a	----------	linking
----------	----------	vowel deletion
doomas |||| σσ σ ------ oo / ||/ ------ \ o |	gari ||| σ σ ---- o / |/ ---- o |	stress assignment
d m s | | | CVVCVCV | | o a	----	V epenthesis
d m s | | | CVVCVCV | | o a	----	linking

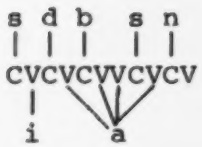
-------	----	C epenthesis
σσ σ σ /||/|/| CVVCVCV ||||||| doomasa	----	resyllabification
do'omasa	'gari	surface form

The following derivation is an example of multiple prefixation.

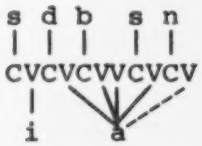
paasan	underlying form
σσ σ /||/| CVVCVC |||||| paasan	syllabification
[[g][paasan]]	word formation
[[g][baasan]]	voicing
----------	vowel deletion
[[dV][g][baasan]]	word formation
----------	voicing
d gb s n | | | | CVCCVVCVC \ \/ a	linking
σ σσ σ /| /||/| CVCCVVCVC ||||||||| dagbaasan	resyllabification
--------	vowel deletion
[[si][dagbaasan]]	word formation
----------	voicing
----------	vowel deletion
σ σ σσ σ /| /| /||/| CV CVCCVVCVC || ||||||||| [[si][dagbaasan]]	resyllabification



stress assignment

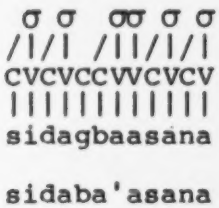


V epenthesis



linking

C epenthesis



resyllabification

sidaba'asana — surface form

The unlinked C (g-) is not realized phonetically.

The following examples demonstrate the C epenthesis rule, which attempts to form a CV syllable between vowels. Between two identical vowels it inserts a glottal stop in emphatic speech. Between different vowels it attempts to link the C to the first segment.

geena	uak	underlying form
σσ σ /||/| CVVCV ||||| geena	σσ || VVC ||| uak	syllabification
-----	-----	voicing
-----	-----	vowel deletion
geena \\\||\\\| σσ σ ----- oo / ||/ ----- \ o \\\|	uak || σσ ----- o / |/ ----- o |	stress assignment
-----	k | VVCV || ua	V epenthesis
-----	k | VVCV | / ua	linking
-----	k | VCVCV | |/ u a	C epenthesis

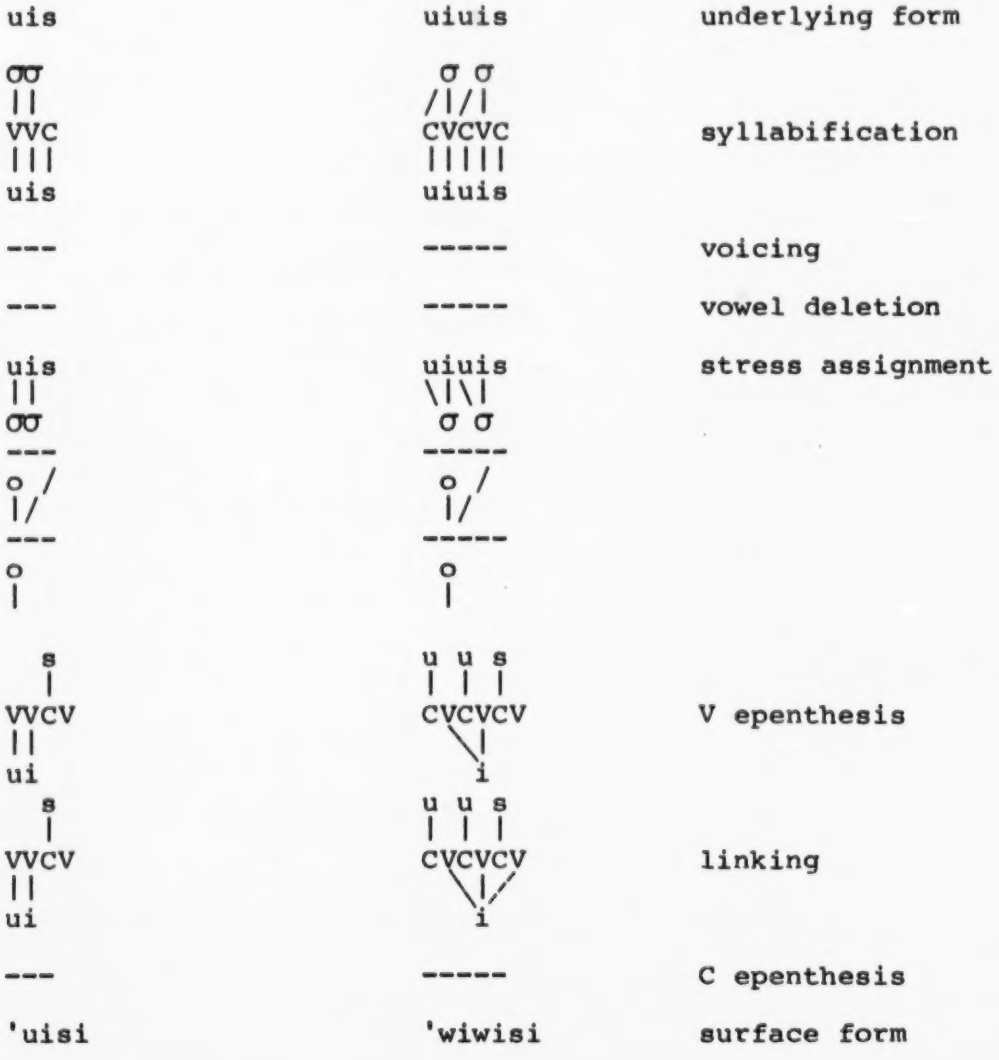
```
                    k
                    |
-----               VCVCV          linking
                    | |
                    u a

ge'ena              'uwaka         surface form
```

The following derivations show how the minimal pairs with respect to stress, ['caana] 'thousand' and [ca'ana] 'pants', are accounted for.

caana	caan	underlying form
σσ σ /||/| CVVCV ||||| caana	σσ /|| CVVC |||| caan	syllabification
-----	----	voicing
-----	----	vowel deletion
caana \\\||\\\| σσ σ ----- σσ / ||/ ----- \ σ \\\|	caan \\\|| σσ ---- σ / |/ ---- σ |	stress assignment
-----	c n | | CVVCV \\/ a	V epenthesis
-----	c n | | CVVCV \\/ a	linking
-----	-----	C epenthesis
ca'ana	'caana	surface form

The apparent contrast between the [u] and [w] in 'uisi, 'flow' and 'wiwisi 'suck' is examined below. Note that within the syllabification of the underlying form that /uis/ is syllabified as two syllables. As was mentioned above, major lexical categories must have at least two syllables in Tabaru.



7 Conclusion

It is clear that lexical phonology handles the facts of the Tabaru language. Indeed, it eliminates the problem of some rules applying in suffixation but not in prefixation, a problem not solved by generative phonology. The predictability of semivowels from the theory of syllable structure is quite a step forward as compared to the previously unclear status of these segments. The notion of closed syllables in the underlying form allows for simplicity in defining stress assignments and predicting reduplication patterns. In other words, we have made considerable progress in the understanding of the phonology of the Tabaru language.

NOTES

[1]What is meant by C and V in this context is a variable number of Cs or Vs.

[2]There are exceptions to the above description of the prefix g-. There are vowel initial roots which appear with an ng- or an n- rather than a g- in their surface forms. The majority of the exceptions take the ng- prefix and there are one or two which take an n-. These cases will be lexically marked. Van der Veen (1915:106), who along with Kern posited an ng- prefix for the North Halmaheran languages, was at the same time suspicious that on vowel initial roots the ng- almost always appeared as a g-.

[3]There are (at least) three rules applying to these forms: metathesis, vowel deletion and deletion of a 3rd person subject prefix. The reflexive prefix ma- is not affected by these rules and is thus not in the same class as the pronominal prefixes.

First, the process of metathesis occurs within the paradigm of the pronominal verb prefixes. This process functions as the following rule defines:

```
   1    2    3    4
   C    V    C    V     -->     3 2 1 4
[+nasal] [+labial]
```

For example:

[[wo][mi]] --> [[mo][wi]]

Second, the vowel deletion rule which applies to the pronominal prefix deletes vowels in the opposite direction as compared to the vowel deletion rule applying to suffixation, for example:

[[nio][a]] --> [[ni][a]]

Third, there is a deletion of the subject prefix when both subject and object are in the third person singular as seen in the following examples (Fortgens 1928:340):

Expected forms:	Actual forms:	Meaning:
wo-wi-paasana	wi-paasana	'he beats him'
mo-mi-paasana	mi-paasana	'she beats her'
mo-wi-paasana	mi-paasana	'she beats him'

The third example is more complex than the first two because the process of metathesis appears to have taken place before deletion.

[4]In Tabaru it is irrelevant whether stratum one or stratum two is first in the derivation.

APPENDICES

Appendix A

LIST OF ABBREVIATIONS

1	first person
2	second person
3	third person
ALR	already
AP1	applicative 1
AP2	applicative 2
AP2	applicative 3
ART	article
CON	continuative
EXE	exclamation
f	feminine singular
HAB	habitual
LOC	locative
m	masculine singular
n	neuter singular or plural
NEG	negative
NML	nominalizer
p	plural
pa	plural all
pe	plural exclusive
pi	plural inclusive
POS	possessive
PO	pronominal object
pp	plural partitive
PROG	progressive
PS	pronominal subject
PST	past
RDP	reduplication
REC	reciprocal
REF	reflexive
REP	repetitive
s	singular
SU	surrounding

Appendix B

INTERLINEAR TARARU TEXT

```
O     haiwani   moi   ma    ronga   o     sarigala.
o     haiuani   moi   ma    ronga   o     sarigala
ART   animal    one   POS   name    ART   fox

Ma    moi   geena   itagitagi,   itagitagi,
ma    moi   geena   i-RDP-tagi   i-RDP-tagi
ART   one   that    it-REP-go    it-REP-go

de    imakamake          de    o     gosomanga.
de    i-   ma- ka- make   de    o     gosomang
and   they-RFL-REC-see   and   ART   crocodile

De    o     babarenoka   o     gosomanga   geena   ingose
de    o     babaren-oka  o     gosomang    geena   i -ngose
and   ART   beach-  LOC  ART   crocodile   that    it-say

"O     ngoi   neena   o     inomo   maoa   tomadamake!
 o     ngoi   neena   o     inom    maoa   to-ma- dV-make
 ART   1S     this    ART   food    good   I- RFL-in-see

O     akeme   isakisaki!"   De    ma    sarigala   geena   ingose
o     akem    i-RDP-saki    de    ma    sarigala   geena   i- ngose
ART   meat    it-REP-fat    and   ART   fox        that    it-say

"O    neena   ka    o     naokowa     kaiamodi.
 o    neena   ka    o     naoko-uua   kaiamodi
 ART  this    not   ART   fish- NEG   is.not

Neena   o     gota   ma    ngogu   dadi   uwa   nagoli!
neena   o     gota   ma    ngogu   dadi   uua   no- a- goli
this    ART   wood   POS   piece   so     NEG   you-it-bite

Neena   ka    o     nyawa,   ka    o     titi.
neena   ka    o     nyaua    ka    o     titi
this    not   ART   person   not   ART   pig

Ka    o     siodea,   dadi   uwa   nagoli!
ka    o     siodea    dadi   uua   no- a- goli
not   ART   pork      so     NEG   you-it-bite

Napidili          ka     geena   ka    o     gota   ma    ngogu!"
no- a- pidil      ka     geena   ka    o     gota   ma    ngogu
you-it-release    only   that    not   ART   wood   POS   piece

So   ma   gosomanga   geena   yapidilokau.
so   ma   gosomang    geena   io-a- pidil-  oka-ou
so   ART  crocodile   that    it-it-release-PST-ALR
```

```
Dadi  ma   sarigala  geena  yoarokau.
dadi  ma   sarigala  geena  io-aro-oka-ou
so    ART  fox       that   it-run-PST-PER

Yoaro   ka    ahade,  ma   sarigala   geena  ingose,
io-aro  ka    ahade   ma   sarigala   geena  i- ngose
it-run  only  then    ART  fox        that   it-say

"E     alele,  alele,   ka    noiodomuwa,
 e     alel    alel     ka    no- i- odom-uua
 EXC   thanks  thanks   only  you-me-eat- NEG

ka    noiodomuwa!"
ka    no- i- odom-uua
only  you-me-eat-NEG

Dadi  ma   gosomanga  geena  yatootasuku.
dadi  ma   gosomang   geena  io-a- tootasuku
so    ART  crocodile  that   it-it-angry

Yatootasuku       ho,   ma   manai   yaasoko
io-a- tootasuku   ho    ma   manai   io-a-   asok
it-it-angry       then  but  friend  it-them-call

imakaasoko.
i-   ma- ka- asok
they-RFL-REC-call

Dadi  manga  gudaioka   yoboa        yoboa         geena
dadi  manga  gudai-oka  io- boa      yoboa         geena
so    they   many- PST  they-arrive  they-arrive  that

yomau       o    serigala  geena  yaodomo,      yatagoko
io- mau     o    serigala  geena  io- a- odom   io-a-tagok
they-want   ART  fox       that   they-it-eat   they-it-catch

yatagoko       duru.
io-a-tagok     duru
they-it-catch  first

O    serigala  geena  ingose,
o    sarigala  geena  i- ngose
ART  fox       that   it-say

"Wasi  tinidoimikasi
 uasi  ti-ni- doimi- oka-osi
 wait  I- you-count- PST-CON

ma   dodogumoka  ahade  niodomo      ngoi."
ma   dodogumoka  ahade  ni- i- odom  ngoi
but  last        then   you-me-eat   me
```

Iɖuanga	ɖe,	o	gosomanga	geena	imahatohato.
iɖuanga	ɖe	o	gosomang	geena	i- ma- RDP-hato
after	and	ART	crocodile	that	they-RFL-REP-row

De	ma	sarigala	geena	iɖoimi,
ɖe	ma	sarigala	geena	i- ɖoim
and	ART	fox	that	it-count

"Moi,	modidi,	saange,	soata,	motoa."
moi	modidi	saange	soat	motoa
one	two	three	four	five

Momongo	ka	yoaroka.	Dadi	o	gosomanga
momong	ka	io-ara-oka	ɖadi	o	gosomanga
in.fact	only	it-run-PST	so	ART	crocodile

geena	yagoli	o	sarigala	ka	iɖadiou.
geena	i-a-goli	o	sarigala	ka	i- ɖadi-ou
that	it-it-bite	ART	fox	not	it-can- ALR

Ipade	o	sarigala	ipade	masala.
i- pade	o	sarigala	i- pade	masala
it-clever	ART	fox	it-clever	very

Free translation:

Once upon a time, there was a fox. As he was walking he met a crocodile at the beach. The crocodile said, "Hey! I've found some good food for myself! This is fat meat!" The fox said, "This is not fish, this is a piece of wood, so don't bite it! This is not a person, not a pig and it's not pork so don't bite it!"

So the crocodile let him go, and the fox ran away. But the fox said, "Thank you! thank you! you didn't eat me! you didn't eat me!" So the crocodile was angry at him. The crocodile was angry so he called his friends. So many of them arrived. They arrived and they wanted to eat the fox. First they caught it. The fox said, "Wait, I'll count you, then you eat me." Then the crocodiles put themselves in a row and the fox counted, "One, two, three, four, five." Then he ran so those crocodiles couldn't bite him. He is clever, that fox, very clever.

REFERENCES

Allison, A.W., H. Barrows, C. Blake, A. Carr, A. Eastman, H. English eds. 1975. The Norton Anthology of Poetry. New York: W.W. Norton and Company.

Bloomfield, L. 1933. Language. Chicago: Holt, Rinehart and Winston.

Chomsky, N. and M. Halle. 1968. The Sound Pattern of English. New York: Harper and Row.

Clements, G.N. and S.J. Keyser. 1983. CV Phonology. Cambridge: MIT Press.

Fortgens, J. 1928. Aanteekeningen van het Tabaroesch, Tabaroesche Volksverhalen en Raadsels. Bijdragen tot de Taal-, Land- en Volkenkunde van Nederlandsch-Indie 84, 300-544. 's-Gravenhage: Nijhoff.

Gleason, H.A. 1961. An Introduction to Descriptive Linguistics. Chicago: Holt, Rinehart and Winston.

Goldsmith, J.A. 1976. Autosegmental Phonology. Bloomington: Indiana University Linguistics Club.

Halle, M. and J.R. Vergnaud. 1982. On the Framework of Autosegmental Phonology. The Structure of Phonological Representations (Part I), ed. by H. van der Hulst and N. Smith, 65-82. Dordtrecht: Foris Publications.

Hammond, M. 1986. The Obligatory-branching Parameter in Metrical Theory. Natural Language and Linguistic Theory, vol. 4. no 2: 185-228. Dordtrecht: D. Reidel Publishing Company.

Hansen, K.C. and L.E. Hansen. 1969. Pintupi Phonology. Oceanic Linguistics vol. 8, 153-170. Honolulu: University of Hawaii Press.

Harris, J.W. 1983. Syllable Structure and Stress in Spanish: A Nonlinear Analysis. Cambridge, Mass: MIT Press.

Hueting, A. 1908. Tobeloreesch-Hollandsch woordenboek met Hollandsch-Tobeloreesch inhoudsopgave. 's-Gravenhage: Nijhoff.

Kahn, D. 1976. Syllable-based Generalizations in English Phonology. New York: Garland Publishing Co.

Kiparsky, P. 1982. Lexical Morphology and Phonology. Linguistics in the Morning Calm. Seoul: Hanshin Publishing Co.

Liberman, M. and A. Prince. 1977. On Stress and Linguistic Rhythm. Linguistic Inquiry 8, 249-336. Cambridge, Mass: MIT Press.

McCarthy, J.J. and A.S. Prince. To appear. Prosodic Morphology.

Mohanan, K.P. 1986. The Theory of Lexical Phonology. Dordtrecht: D. Reidel Publishing Company.

Perlmutter, D.M. 1978. Impersonal Passives and the Unaccusative Hypothesis. Proceedings of the Sixth Annual Meetings of the Berkeley Linguistic Society. Berkeley: University of California.

Pike, K.L. 1947. Phonemics: a Technique for Reducing Languages to Writing. Ann Arbor: University of Michigan Press.

Pike, K.L. 1967. Language in Relation to a Unified Theory of the Structure of Human Behavior. The Hague: Mouton and Co.

Tuggy, D. 1979. On Angle Brackets. Linguistic Notes from La Jolla no. 7, 106-123. San Diego: University of California.

van Baarda, M.J. 1895. Woordenlijst. Galelareesch-Hallandsch. Met ethnologische aanteekeningen, op de wooden, die daartoe annleiding gaven. s'-Gravenhage: Nijhoff.

van der Veen H. 1915. De Noord-Halmahera'se taalgroep tegenover de Austronesiese talen. Leiden: Van Nifterik.

Voorhoeve, C.L. 1981. The Non-Austronesian Languages in the North Moluccas. Canberra: Australian National University.

Voorhoeve, C.L. 1984. Comparative Linguistics and the West Papuan Phylum. Ambon: Lembaga Ilmu Pengetahuan Indonesia.

Yoshida, S. 1980. Folk Orientation in Halmahera with Special Reference to Insular Southeast Asia. The Galela of Halmahera, ed. by N. Ishige, 19-88. Osaka: National Museum of Ethnology.